

Effect of Heterogeneity on Decorrelation Mechanisms in Spiking Neural Networks: A Neuromorphic-Hardware Study

Thomas Pfeil,^{1,*} Jakob Jordan,² Tom Tetzlaff,² Andreas Grübl,¹ Johannes Schemmel,¹
Markus Diesmann,^{2,3,4} and Karlheinz Meier¹

¹*Kirchhoff-Institute for Physics, Heidelberg University, 69120 Heidelberg, Germany*

²*Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6)
and JARA BRAIN Institute I, Jülich Research Centre, 52425 Jülich, Germany*

³*Department of Psychiatry, Psychotherapy and Psychosomatics, Medical Faculty,
RWTH Aachen University, 52074 Aachen, Germany*

⁴*Department of Physics, Faculty I, RWTH Aachen University, 52074 Aachen, Germany*

(Received 28 November 2014; revised manuscript received 19 February 2016; published 18 May 2016)

High-level brain function, such as memory, classification, or reasoning, can be realized by means of recurrent networks of simplified model neurons. Analog neuromorphic hardware constitutes a fast and energy-efficient substrate for the implementation of such neural computing architectures in technical applications and neuroscientific research. The functional performance of neural networks is often critically dependent on the level of correlations in the neural activity. In finite networks, correlations are typically inevitable due to shared presynaptic input. Recent theoretical studies have shown that inhibitory feedback, abundant in biological neural networks, can actively suppress these shared-input correlations and thereby enable neurons to fire nearly independently. For networks of spiking neurons, the decorrelating effect of inhibitory feedback has so far been explicitly demonstrated only for homogeneous networks of neurons with linear subthreshold dynamics. Theory, however, suggests that the effect is a general phenomenon, present in any system with sufficient inhibitory feedback, irrespective of the details of the network structure or the neuronal and synaptic properties. Here, we investigate the effect of network heterogeneity on correlations in sparse, random networks of inhibitory neurons with nonlinear, conductance-based synapses. Emulations of these networks on the analog neuromorphic-hardware system Spikey allow us to test the efficiency of decorrelation by inhibitory feedback in the presence of hardware-specific heterogeneities. The configurability of the hardware substrate enables us to modulate the extent of heterogeneity in a systematic manner. We selectively study the effects of shared input and recurrent connections on correlations in membrane potentials and spike trains. Our results confirm that shared-input correlations are actively suppressed by inhibitory feedback also in highly heterogeneous networks exhibiting broad, heavy-tailed firing-rate distributions. In line with former studies, cell heterogeneities reduce shared-input correlations. Overall, however, correlations in the recurrent system can increase with the level of heterogeneity as a consequence of diminished effective negative feedback.

DOI: [10.1103/PhysRevX.6.021023](https://doi.org/10.1103/PhysRevX.6.021023)

Subject Areas: Complex Systems,
Computational Physics,
Interdisciplinary Physics

I. INTRODUCTION

Dynamical systems in nature often exhibit a remarkable degree of diversity, specialization, or anticorrelation across their components, despite equalizing factors such as common input or homogeneity in component and interaction parameters. In many cases, these observations

can be explained by the effect of negative feedback. Cell differentiation caused by lateral inhibition [1], formation of new species driven by competition [2], or antiferromagnetism [3] constitute just a few examples. In recurrent neuronal networks, inhibitory feedback constitutes a powerful decorrelation mechanism that allows different neurons to respond nearly independently, even if they are driven by largely overlapping local or external inputs [4–6]. Decorrelation by negative feedback hence implements an efficient form of redundancy reduction. In biological systems, it may serve similar purposes as decorrelation in technical applications, where it is used in data compression (e.g., principal-component analysis [7]), cross-talk reduction (e.g., in digital signal processing [8]), echo suppression

*Corresponding author.

thomas.pfeil@kip.uni-heidelberg.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

(e.g., in acoustics [9]), or random-number generation in hardware [10]. Moreover, inhibitory feedback suppresses “quantization noise” at low frequencies and can thereby increase the dynamical range and signal-to-noise ratio for the encoding of analog signals in the spiking activity of recurrent neural networks [11]. It is tempting to exploit these mechanisms in synthetic, neurally inspired architectures such as analog neuromorphic hardware.

Analog neuromorphic hardware mimics properties of biological neural systems using physical models of neurons and synapses (capacitors, for example, emulate insulating cell membranes) [12,13]. The temporal evolution of the analog circuits represents a solution to the corresponding model equations. In consequence, neural-network emulations on analog neuromorphic hardware are massively parallel, extremely fast, and energy efficient. Analog neuromorphic devices are, therefore, highly attractive as tools for neuroscientific research, e.g., for the investigation of learning on long time scales, and technical applications [14–17]. A biologically inspired neural network (olfactory system of insects) performing rapid online data (odor) classification, for example, has recently been successfully implemented on the analog neuromorphic-hardware system Spikey [18,19]. In this application, decorrelation by inhibition is an essential ingredient to guarantee high classification performance. The suppression of quantization noise by inhibitory feedback [11] has been used as a means of noise shaping in several neuromorphic-hardware applications, aiming at the construction of biologically inspired ultra-low-power analog-to-digital converters [20,21].

For the functional performance of neuronal architectures, the level of correlations between the activities of individual neurons is often pivotal. Whether such correlations are beneficial or not is context dependent. A number of previous studies emphasize a functional benefit of certain types of correlation for encoding (decoding) of information in (from) populations of neurons [22–24], information transmission [25–27], robustness against noise [28], or gain control of postsynaptic neurons [29]. Other studies argue that positive cross-correlations are detrimental as they decrease the precision or sparseness of population codes [19,30–33]. Cohen and Maunsell [34], for example, have shown that decreased spike-train correlations in macaque visual area V4 are accompanied by increased behavioral performance in an orientation change-detection task. Depending on the similarity between the trial-averaged responses of different neurons to external stimuli (signal correlation), noise correlations (correlations not explained by signal correlations) can either increase or decrease the amount of information that can be encoded in or decoded from a population of neurons. In populations of neurons with high signal correlation, vanishing or even negative noise correlations are desirable to improve the population code [23].

In finite neural networks, an inevitable source of correlated neural activity is common presynaptic input, shared by multiple postsynaptic neurons. In network models and *in vivo* recordings, however, pairwise correlations in the activity of neighboring neurons have been found to be substantially smaller than expected given the amount of shared input [4,5,35–38]. In several studies, this observation has been explained by inhibitory coupling. While Ly *et al.* [39] and Middleton *et al.* [40] primarily focused on the effect of feedforward inhibition, Renart *et al.* [4], Wiechert *et al.* [41], and Tetzlaff *et al.* [5] attributed the smallness of correlations to an active decorrelation of neural activity by inhibitory feedback. The mechanism underlying this active decorrelation has already been described by Mar *et al.* [11]. In this study, the authors focused on the suppression of low-frequency fluctuations of the population firing rate by recurrent dynamics. As the amplitude of population-rate fluctuations is directly linked to pairwise correlations (see, e.g., Ref. [42]), the effect described in Ref. [11] corresponds to a suppression of pairwise correlations in the spiking activity. The theory underlying decorrelation by inhibitory feedback suggests the effect to be general: Decorrelation should be observable in any system with sufficiently strong inhibitory feedback, irrespective of the details of the network structure and the cell and synapse properties. For networks of spiking neurons, however, the effect has so far been explicitly demonstrated only for the homogeneous case, where all neurons have identical properties, receive (approximately) the same number of inputs, and, hence, fire at about the same rate [4,5]. Moreover, the subthreshold dynamics of individual neurons was assumed to be linear.

Biological neuronal networks typically exhibit broad, heavy-tailed firing-rate distributions [43–49], indicating a high degree of heterogeneity, e.g., in synaptic weights [50–54], in-degrees [55], or time constants [48,56]. The same holds for neural networks implemented on analog neuromorphic hardware. All analog circuits suffer from device variations caused by unavoidable variability in the manufacturing process. Neurons and synapses implemented in analog neuromorphic hardware therefore exhibit heterogeneous response properties, similar to their biological counterparts [57,58]. To understand the dynamics and function of recurrent neural networks in both biological and synthetic substrates, it is therefore essential to account for such heterogeneities.

Previous work on recurrent neural networks has shown that heterogeneity in single-neuron properties or connectivity broadens the distribution of firing rates [48,59] and affects the stability of asynchronous or oscillatory states [55,60–64]. A number of studies pointed at a potential benefit of heterogeneity for the information-processing capabilities of neural networks [64–75]. The effect of heterogeneity on correlations in the activity of recurrent networks of spiking neurons, however, remains unclear.

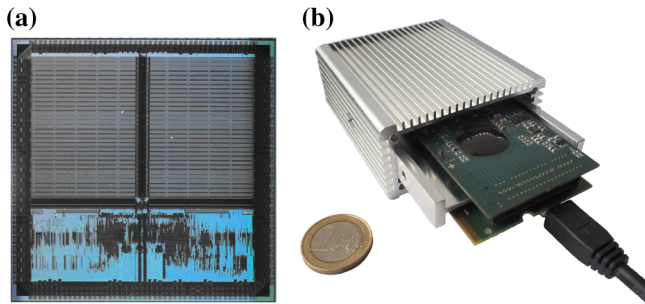


FIG. 1. The neuromorphic-hardware system Spikey. (a) Photograph of the Spikey chip (size $5 \times 5 \text{ mm}^2$). It comprises analog circuits of 384 neurons and 98304 synapses, is highly configurable, and emulates neural-network dynamics with a speed-up of 10^4 with respect to biological real time. (b) Photograph of the partly cased Spikey system, carrying the Spikey chip (covered by a black round seal) and conventional memory. The system is connected to the host computer via USB 2.0, consumes 6 W of power in total and less than 1 nJ per synaptic transmission (see Supplemental Material, Sec. I [78]).

Padmanabhan and Urban [69] have shown that the responses of a population of unconnected neurons are decorrelated by heterogeneity in the neuronal response properties. These results are supported by the subsequent theoretical analysis in Ref. [72]. In the following, we refer to this type of decorrelation by heterogeneity as feedforward decorrelation. It does not account for the effect of the recurrent-network dynamics. Active decorrelation due to inhibitory feedback (see above and Refs. [4,5]), in contrast, constitutes a very different mechanism. The effect of heterogeneity on this feedback decorrelation has recently been studied by Bernacchia and Wang [76] in the framework of a recurrent network of linear firing-rate neurons. In this setup, correlations are suppressed by heterogeneity in the network connectivity (distributions of coupling strengths or random dilution of connectivity). It remains unclear, however, whether this holds true for networks of (nonlinear) spiking neurons.

In this study, we investigate the impact of heterogeneity on input and output correlations in the asynchronous regime of sparse networks of leaky integrate-and-fire (LIF) neurons with conductance-based synapses. Emulation of the networks on the analog neuromorphic-hardware system Spikey (Fig. 1) [18,77] enables us to investigate the impact of substrate-specific properties on the network dynamics. Insights about the interplay between features of the computing substrate and network dynamics are a necessary prerequisite for the development of algorithms that exploit the benefits of analog neuromorphic systems at best.

The configurability of this system [18] enables us to systematically vary the level of heterogeneity, and to disentangle the effects of heterogeneity on feedforward and feedback decorrelation (see above). For simplicity, we focus on purely inhibitory networks, thereby emphasizing

that active decorrelation by inhibitory feedback does not rely on a dynamical balance between excitation and inhibition [5,6]. We show that decorrelation by inhibitory feedback is effective even in highly heterogeneous networks with broad distributions of firing rates (Sec. III A). Increasing the level of heterogeneity has two effects: Feedforward decorrelation is enhanced and feedback decorrelation is impaired. Because of the latter, the overall input and output correlations do not necessarily become smaller with increasing heterogeneity. They can even increase (Sec. III B).

Note that results from specific network emulations on hardware do not directly translate to those obtained by simulations on conventional computers, because the dynamics, parametrization, and interplay of analog circuits are very complex and difficult to reproduce with classical simulations. If simplified models for spatial and temporal variability are considered in software simulations, however, emulation results can be reproduced qualitatively, thereby verifying the design of the hardware system. While our hardware system is designed to physically implement biologically realistic neural algorithms in a fast and energy-efficient way, software simulations are used as a complementary tool to isolate, verify, and investigate different hardware features, such as spatial and temporal parameter variations. Because of the limited access and configurability of network parameters, this would be difficult to achieve with hardware studies alone. In analogy to the necessity of performing experiments on biological neural systems to verify assumptions made in computational neuroscience, actual emulations on neuromorphic hardware are essential to understand its properties and develop efficient neural algorithms for these devices. The fact that our main findings hold true for both emulations on hardware and simulations with software, and that they can be distilled to simple linear models, supports their broad relevance and robustness.

II. METHODS

A. Network model

Details on the network, neuron, and synapse model are provided in Table I. Parameter values are given in Table II. Briefly, we consider a purely inhibitory, sparse network of N ($N = 192$, unless stated otherwise) LIF neurons with conductance-based synapses. Each neuron receives input from a fixed number $K = 15$ of randomly chosen presynaptic sources, independently of the network size N . Self-connections and multiple connections between neurons are excluded. Resting potentials E_1 are set above the firing thresholds Θ (equivalent to applying a constant supra-threshold input current). We thereby ensure autonomous firing in the absence of any further external input. Because of temporal noise, the initial conditions are essentially random.

B. Network emulations on the neuromorphic-hardware system Spikey

The Spikey chip (Fig. 1) consists of physical models of LIF neurons and conductance-based synapses with exponentially decaying dynamics (for details, see Table I). The emergent dynamics of these physical models represents a solution for the model equations of neurons and synapses in continuous time, in parallel for all units. In contrast, in classical simulations on von Neumann architectures, model equations are solved by stepwise numerical integration, where parallelization is limited by the available number of processor cores. To emphasize the difference between simulations using software and simulations using physical models, the term emulation is used for the latter [18].

The response properties of physical neurons and synapses vary across the chip due to unavoidable variations in the production process that manifest in a spatially disordered pattern (fixed-pattern noise). In contrast to the approximately static fixed-pattern noise, temporal noise, including electronic noise and transient experiment conditions (e.g., chip temperature), impairs the reproducibility of emulations. In general, two network emulations with identical configuration and stimulation do not result in identical network activity. Both fixed-pattern and temporal noise need to be taken into account when developing models for analog neuromorphic hardware.

The key features of the Spikey chip are the high acceleration and configurability of the analog network implementation. Some network parameters, e.g., synaptic weights and leak conductances, are configurable for each unit, while other parameters are shared for several units (for details, see Ref. [18]). The hardware system is optimized for spike input and output and allows us to record the membrane potential of one (arbitrarily chosen) neuron with a sampling frequency of 96 MHz in hardware time. On the Spikey chip, capacitances are smaller and conductances are much higher than in biological nervous

systems. In consequence, networks on the Spikey chip are emulated with a speed-up of approximately 10^4 with respect to biological real time. Because of this high acceleration of the neuromorphic chip, the data bandwidth of the connection between the neuromorphic system and the host computer is not sufficient to communicate with the chip in real time. Consequently, input and output spikes (for stimulation and from recordings, respectively) are buffered in a local memory next to the chip. The high acceleration of the Spikey chip allows most of the transistors to operate outside of weak inversion, thereby reducing the effect of transistor variations and minimizing fixed-pattern noise.

In contrast to such accelerated systems, most other configurable, analog neuromorphic substrates are designed for real-time emulations at very low power consumption [79–85] and implement fewer, but more complex, neurons [86,87].

Access to the Spikey system is encapsulated by the simulator-independent language PyNN [88,89], providing a stable and user-friendly interface. PyNN integrates the hardware into the computational neuroscience tool chain and has facilitated the implementation of several network models on the Spikey chip [18,19,90–92].

On the Spikey system, a spiking neural network is emulated as follows [Fig. 2(a)]. First, the network described in PyNN is mapped to the Spikey chip, i.e., neurons and synapses are allocated and parametrized. Second, input spikes, if available, are prepared on the host computer and transferred to the local memory on the hardware system. Third, the emulation is triggered and available input spikes are generated. Output spikes and membrane data are recorded to local memory. Last, spike and membrane data are transferred to the host computer and scaled back into the biological domain of the PyNN model description.

For consistency with the model description and simplified comparison to the existing literature, all hardware times and all hardware voltages are expressed in terms of

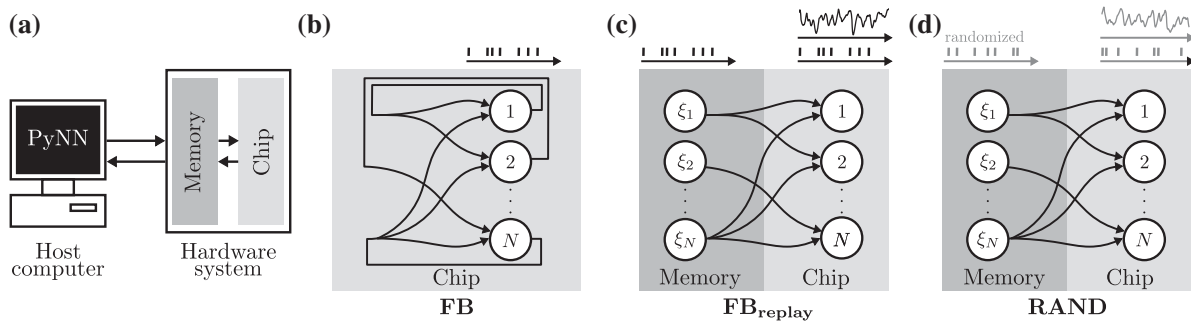


FIG. 2. Experimental setup. (a) Data flow of the Spikey system. For details, see Sec. II B. (b) Network with on-chip feedback connections (FB). Spikes from all neurons are recorded to the local memory. (c) Spikes of the FB network in (b) replayed from memory via off-chip spike sources ξ_i to neurons i (FB_{replay}). Spike times of ξ_i correspond to those recorded from neuron i in (b). Spikes from all neurons or the free membrane potential of one selected neuron are recorded. (d) Like (c), but spike times from (b) are randomized for each source ξ_i (RAND).

the quantities they represent in the neurobiological model throughout this study.

C. Experimental setup

To differentiate and compare the effects of shared inputs and feedback connections on correlations, we investigate two different emulation scenarios: First, we emulate networks with intact feedback (FB) [Fig. 2(b)], and, second, the contribution of shared input is isolated by randomizing the temporal order of this feedback (RAND) [Fig. 2(d)].

In the RAND scenario, the inputs of neurons are decoupled from their outputs. Spatiotemporal correlations in presynaptic spike trains are removed by randomizing the presynaptic spike times.

Input correlations between neurons are measured via their free membrane potential, i.e., the membrane potential with disabled spiking mechanism (technically, the threshold is set very high). Because membrane-potential traces can be recorded in the hardware only one at a time, traces are obtained consecutively, while repeatedly replaying the previously recorded activity of the FB network to a population of unconnected neurons of equal size. We keep the connectivity the same, and, hence, each neuron receives the same number of spikes as in the recurrent network during the whole emulation, either without (FB_{replay}) [Fig. 2(c)] or with randomization of presynaptic spike times (RAND) [Fig. 2(d)], respectively. To preserve the fixed pattern of variability of synaptic weights in hardware, the same hardware synapses are used for each connection in both scenarios. If network dynamics were reproduced perfectly, membrane-potential traces and spike times would be identical in the FB and FB_{replay} cases (see also Sec. II D).

Drawing two different network realizations (i.e., the connectivity matrix) results in the allocation of different hardware synapses, and, due to fixed-pattern noise, in different values of synaptic weights. To average over this variability, throughout this study, emulation results are averaged over $M = 100$ network realizations, if not stated otherwise.

D. Reproducibility of hardware emulations

Since the initial conditions of the recurrent network on hardware are undefined, consecutive emulations of the FB network result in different network activities. In the RAND and FB_{replay} case, however, the input of neurons is decoupled from their output. Although unavoidable temporal noise is present, the system's state space trajectory returns to the trajectory of the previously recorded FB case. A certain degree of reproducibility is required for two reasons: First, the investigated effect of decorrelation by inhibitory feedback requires a precise relation between spike input and output. Thus, our method of replacing the feedback loop by replay is valid only if temporal noise does not substantially corrupt this relationship. Second, to record the membrane potentials of all neurons, as if recorded at

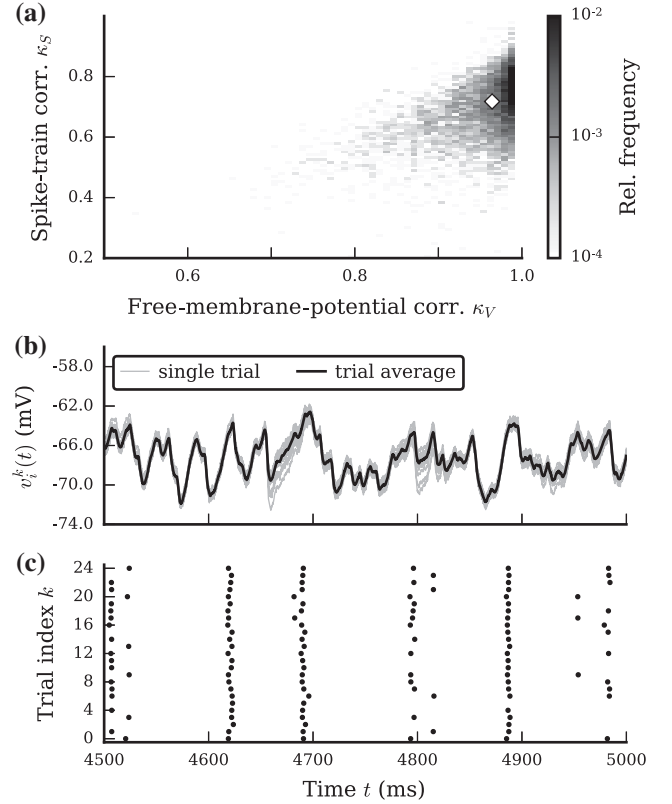


FIG. 3. Reproducibility of free membrane potentials and spiking activity in the FB_{replay} case. (a) Low-frequency coherence κ_V and κ_S of free membrane potentials $v_i^k(t)$ and binned spike trains $s_i^k(t)$ and $s_i^l(t)$, respectively, for each neuron i averaged over $L = 25$ trials k, l , with $k \neq l$, for $M = 50$ different network realizations. The diamond marks the average across all neurons i and M network realizations ($\kappa_V = 0.96$, $\kappa_S = 0.72$). (b) Free single-trial membrane potentials $v_i^k(t)$ (gray lines) and average over trials $(1/L) \sum_{k=1}^L v_i^k(t)$ (black line), and (c) spike density $\xi_i(t)$ of a single neuron i for $L = 25$ identical trials. The selected neuron i has membrane-potential coherence and spike-train coherence closest to the diamond in (a).

once, neuron dynamics have to be reasonably similar in consecutive emulations.

We measure the reproducibility of neuron dynamics by comparing consecutive emulations with identical configuration, i.e., connectivity and stimulation. For this purpose the spiking activity of a FB network is first recorded [Fig. 2(b)] and then repeatedly replayed [Fig. 2(c)]. Reproducibility is quantified by the correlations (κ_X in Table III) of free-membrane-potential traces and output spike trains obtained for individual neurons in $L = 25$ different trials.

Free membrane potentials are reproduced quite well, while spike trains show larger deviations across trials (Fig. 3). Small deviations in the membrane potential [Fig. 3(b)] are amplified by the thresholding procedure [41,93,94] and can lead to large differences between spike trains [Fig. 3(c)]. Consequently, measures based on data of several consecutive replays are more precise for membrane

potentials than for spike trains. Nevertheless, results have to be interpreted with care in both cases.

E. Calibration

The heterogeneity of the Spikey hardware is adjusted by calibrating the leak conductance (since capacitances and potentials cannot be configured individually for each hardware neuron [18]) for each individual neuron, compensating for fixed-pattern noise of neuron parameters. To this end, a population of unconnected neurons is driven by a constant suprathreshold current and the time-averaged population activity \bar{r} is measured. Then, we apply the bisection method [95] to adjust the leak conductance g_l of each neuron, such that the neuron's firing rate matches the target rate \bar{r} . This results in calibration values b for the leak conductance $g_l = g_{l,0}(1 + b)$, where $g_{l,0}$ is the leak conductance before calibration. Because emulations on hardware are not perfectly reproducible, more precise calibration is achieved by evaluating the median over 25 identically configured trials instead of single trials. Furthermore, the bisection method is modified for noisy systems (for details, see Supplemental Material, Sec. II [78]).

Intermediate calibration states are obtained by linearly scaling the full calibration:

$$g_l = g_{l,0}[1 + (1 - a)b]. \quad (1)$$

The heterogeneity a is chosen in $[0, 1]$ for calibrations between the uncalibrated ($a = 1$) and calibrated state ($a = 0$). In the following, the fully calibrated chip ($a = 0$) is used, if not stated otherwise.

This calibration substantially narrows the distribution of firing rates compared to the uncalibrated state (Fig. 4). With respect to the stationary firing rate, variability on the neuron level is reduced from 35.1 to 0.9 s⁻¹.

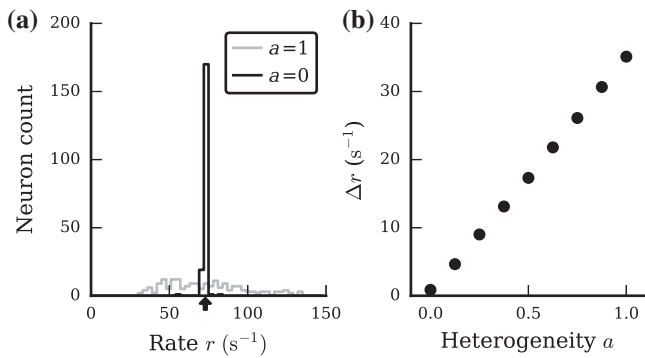


FIG. 4. Calibration of the Spikey chip. (a) Histogram of firing rates r for a population of unconnected neurons with suprathreshold input currents, before (gray line) and after (black line) calibration, each neuron averaged over $L = 100$ trials. The arrow denotes the target rate \bar{r} . (b) Difference $\Delta r = r_{P75} - r_{P25}$ of 75th and 25th percentile of the histograms in (a), as a function of network heterogeneity a [Eq. (1)]. The mean firing rate over all values of a is (73.4 ± 0.3) s⁻¹.

Even in the fully calibrated state, leak conductances can still be widely distributed. Because of the chosen calibration procedure, they are likely to be correlated to other parameters that influence the neurons' response to a constant suprathreshold current after calibration. This mutual compensation can lead to similar phenomenology (here, firing rates) despite disparate parameter values, similar to what is observed in biology [96]. In addition to the remaining variations in neuron parameters, synaptic parameters are significantly distributed [19,97].

F. Correlation measures

In the following, we introduce definitions used to analyze the recorded data. For clarity, all relevant equations and their parametrization are listed in Tables III and IV, respectively.

We quantify correlations of membrane potentials $v_i(t)$ and spike trains $s_i(t)$ by the population-averaged low-frequency coherence κ_V and κ_S , respectively. At frequency zero, the coherence corresponds to the normalized integral of the cross-covariance function, i.e., it measures correlations on all time scales. We define the low-frequency coherence κ_X , with $X \in \{S, V\}$, to be the average coherence over a frequency interval from 0.1 to 20 Hz. In this interval, the suppression of population-rate fluctuations in recurrent networks due to inhibitory feedback is most pronounced, and the coherence is approximately constant. Before calculating the coherence, we convolve the power and cross spectra with a rectangular window to average out random fluctuations. This measure, or a variant of it, is commonly used in the neuroscientific literature [4,5,72,94,98–100]. We use the terms low-frequency coherence and correlation interchangeably.

Throughout this study, the term input correlations is used for correlations between free membrane potentials, and the term output correlations is used for correlations between spike trains. Shared-input correlations are membrane-potential correlations that are exclusively caused by overlapping presynaptic sources, ignoring possible correlations in the presynaptic activity. In homogeneous networks, the average pairwise shared-input correlation

$$\kappa_V = \frac{K}{N} \quad (2)$$

is given by the connectivity K/N [5]. In heterogeneous networks, shared-input correlations can be reduced. In the presence of heterogeneous synaptic weights, for example, the shared-input correlation

$$\kappa_V = \frac{1}{1 + CV_J^2} \frac{K}{N} \quad (3)$$

is decreased by a factor of $1/(1 + CV_J^2)$, where CV_J denotes the coefficient of variation of the (nonzero)

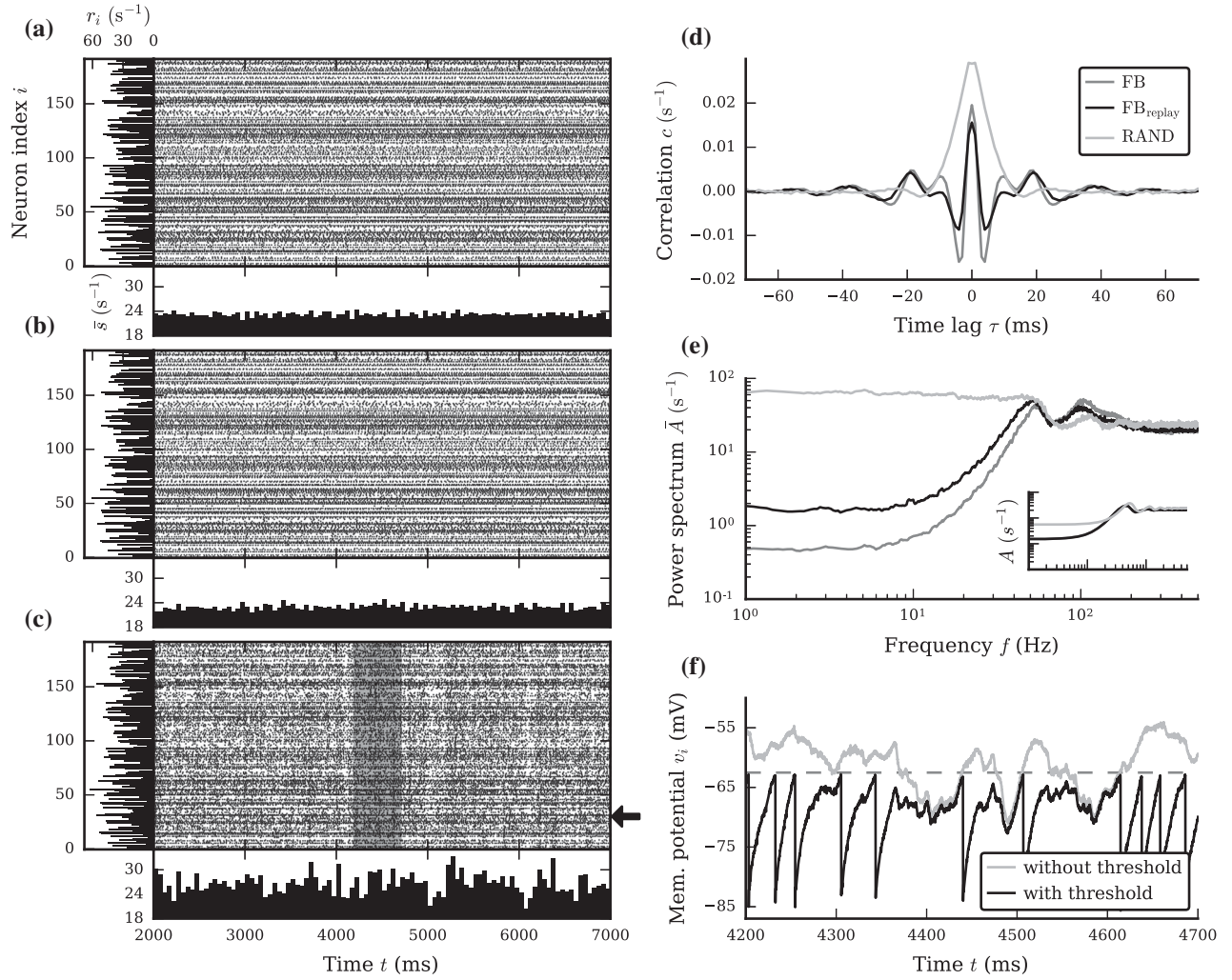


FIG. 5. Typical spiking and membrane-potential activity of a random inhibitory network of LIF neurons with intact and cut feedback loop emulated on the fully calibrated system. (a)–(c) Spiking activity (raster plots), population activity $\bar{s}(t)$ (horizontal histograms, bin size 50 ms), and time-averaged single-neuron firing rates r_i (vertical histograms) in the network with intact feedback (a) and for cases where the feedback loop is cut (b),(c). (a) Intact recurrent network (FB scenario). (b) Population of mutually unconnected neurons receiving input spike trains identical to those in (a) (FB_{replay} scenario). (c) As in (b), but after randomization of presynaptic spike times (RAND scenario). (d),(e) Population-averaged cross-correlation functions $c(\tau)$ (after offset subtraction) of pairs of spike trains (d) and power spectra $\bar{A}(f)$ [(e) log-log representation] of the population activity $\bar{s}(t)$ [cf. horizontal histograms in (a)–(c)] for the FB (dark gray line), FB_{replay} (black line), and RAND scenario (light gray line). Inset in (e): Population-averaged power spectra $A(f)$ of individual single-cell spike trains (same scales as in main panel). Correlation functions and spectra are averaged across $M = 100$ network realizations. (f) Membrane potential of a neuron in the RAND scenario [with firing rate of 23.20 s^{-1} close to population average of 23.24 s^{-1} ; see black arrow in (c)] with intact (black curve) and removed threshold (gray curve; free membrane potential). The threshold potential is marked by the horizontal dashed line. The time frame corresponds to the gray shaded region in (c).

synaptic weights. Note, however, that heterogeneities which affect only the spike generation but not the integration of synaptic inputs, e.g., distributions of firing thresholds, have no effect on the shared-input correlation.

We assess the significance of correlations by comparing the results from emulations to correlations in surrogate data, in which we remove spatial correlations. For every neuron, we randomly shuffle bins of the membrane-potential trace and assign a new time stamp uniformly drawn from the emulation interval to every spike,

respectively. We thereby remove all spatiotemporal correlations between neurons recorded in parallel. By this procedure, we create 100 surrogate trials, across which we calculate the average correlations and the standard error.

To quantify fluctuations in the population activity \bar{s} [Figs. 5(a)–5(c), horizontal histograms] we compute the power spectrum $\bar{A}(f)$ of the population activity [Fig. 5(e)], which we scale with the duration T of the emulation. Consequently, the population power spectrum $\bar{A}(f)$, scaled by the population size, coincides with the

time-averaged population activity \bar{r} for high frequencies: $\lim_{f \rightarrow \infty} (1/N) \bar{A}(f) = \bar{r}$ [37].

As a measure of pairwise correlations in the time domain [Fig. 5(d)], we compute the population-averaged cross-correlation function $c(\tau)$ by Fourier transforming the population-averaged cross spectrum $C(f)$ to time domain.

III. RESULTS

In this study, we investigate the roles of shared input, feedback, and heterogeneity on input and output correlations in random, sparse networks of inhibitory LIF neurons with conductance-based synapses (Table I), implemented on the analog neuromorphic-hardware chip Spikey (Fig. 1). Similarly to Ref. [5], we separate the contributions of shared input and feedback by studying different network scenarios (Fig. 2): In the FB case, we emulate the recurrent network with intact feedback loop [Fig. 2(b)] and record its spiking activity [Fig. 5(a)]. In the FB_{replay} case [Fig. 2(c)], the feedback loop is cut and replaced by the activity recorded in the FB network. Ideally, the input to each neuron in the FB_{replay} case should be identical to the input of the corresponding neuron in the FB network. As the replay of spikes and the resulting postsynaptic currents and membrane potentials are not perfectly reproducible on the Spikey chip, the neural responses in the FB and in the FB_{replay} scenario are slightly different [compare Figs. 5(a) and 5(b)]. In the RAND case [Figs. 2(d) and 5(c)], we use the same setup as in the FB_{replay} case. However, the spike times in each presynaptic spike train are randomized. While the average presynaptic firing rates and the shared-input structure are exactly preserved in this scenario, the spatio-temporal correlations in the presynaptic spiking activity are destroyed.

Using this setup, we first demonstrate in Sec. III A that active decorrelation by inhibitory feedback [4,5] is effective in heterogeneous networks with conductance-based synapses over a range of different network sizes. In Sec. III B, we show that decreasing the level of heterogeneity by calibration of hardware neurons leads to an enhancement of this active decorrelation.

A. Decorrelation by inhibitory feedback

The time-averaged population activities in the FB, FB_{replay}, and RAND scenarios are roughly identical [vertical histograms in Figs. 5(a)–5(c); see also high-frequency power in Fig. 5(e)]. In the FB and FB_{replay} scenario, fluctuations in the population-averaged activity are small [horizontal histograms in Figs. 5(a) and 5(b)]. The removal of spatial and temporal correlations in the presynaptic spike trains in the RAND case leads to a significant increase in the fluctuations of the population-averaged response activity [horizontal histogram in Fig. 5(c)]. At low frequencies (≤ 20 Hz), the population-rate power in the FB and in the RAND case differs by about 2 orders of magnitude

[dark and light gray curves in Fig. 5(e)]. This increase in low-frequency fluctuations in the RAND case is mainly caused by an increase in pairwise correlations in the spiking activity [Fig. 5(d); the power spectra of individual spike trains [inset in Fig. 5(e)] are only marginally affected by a randomization of presynaptic spike times] [5]. In other words, shared-input correlations, i.e., those leading to large spike-train correlations in the RAND scenario, are efficiently suppressed by the feedback loop in the FB case.

On the neuromorphic hardware, the replay of network activity is not perfectly reproducible (Sec. II D). While the across-trial variability in membrane potentials is small, postsynaptic spikes are dithered by a few milliseconds (Fig. 3). In the FB_{replay} case, the suppression of shared-input correlations by correlations in presynaptic spike trains is slightly less efficient as compared to the intact network (FB). The differences in the population-rate power spectra and in the spike-train correlations between the FB_{replay} and RAND case, respectively, are nevertheless substantial [solid black and light gray curves in Figs. 5(d) and 5(e); note the logarithmic scale; for a detailed investigation of spike dither, see Supplemental Material Sec. IV and Fig. 7 [78]].

Note that the suppression of correlations and, hence, population-rate fluctuations by inhibitory feedback is restricted to low frequencies [here, to frequencies < 50 Hz; see Fig. 5(e)]. In the remainder of this study, we quantify pairwise correlations by the low-frequency coherence in the range 0.1–20 Hz (see Sec. II F). At higher frequencies, the population-rate power spectra in the FB, FB_{replay}, and RAND cases are similar. In Fig. 5(e), the peaks at ~ 50 Hz and higher harmonics result from the single-cell spike-train statistics (they are also visible in the single-cell spectra; see inset): A large fraction of cells, in particular those firing at higher rates, generate regular spike trains with low interspike-interval (ISI) variability [cf. Figs. 6(d) and 6(f)]. These (fast spiking) cells contribute maxima to the spike-train spectra at frequencies close to their firing rates (and higher harmonics). The structure of the population-rate spectra at higher frequencies (≥ 50 Hz) is reproduced using surrogate data where the ISI distributions of the individual neurons (and, hence, their firing rates and ISI variability) are preserved, but serial ISI correlations and cross-correlations between spike trains are destroyed (data not shown).

In the RAND case, presynaptic spike-train correlations are removed, and, hence, input (i.e., free-membrane-potential) correlations are exclusively determined by the number of shared presynaptic sources [Eq. (2)]. If the in-degree K is fixed, input correlations will decrease with network size N [Eq. (2), light gray curve and symbols in Fig. 7(a)]. In purely inhibitory networks with intact feedback loop (FB scenario), correlations in presynaptic spike trains are, on average, significantly smaller than zero [dark gray diamonds in Fig. 7(b)] [5], and largely cancel the

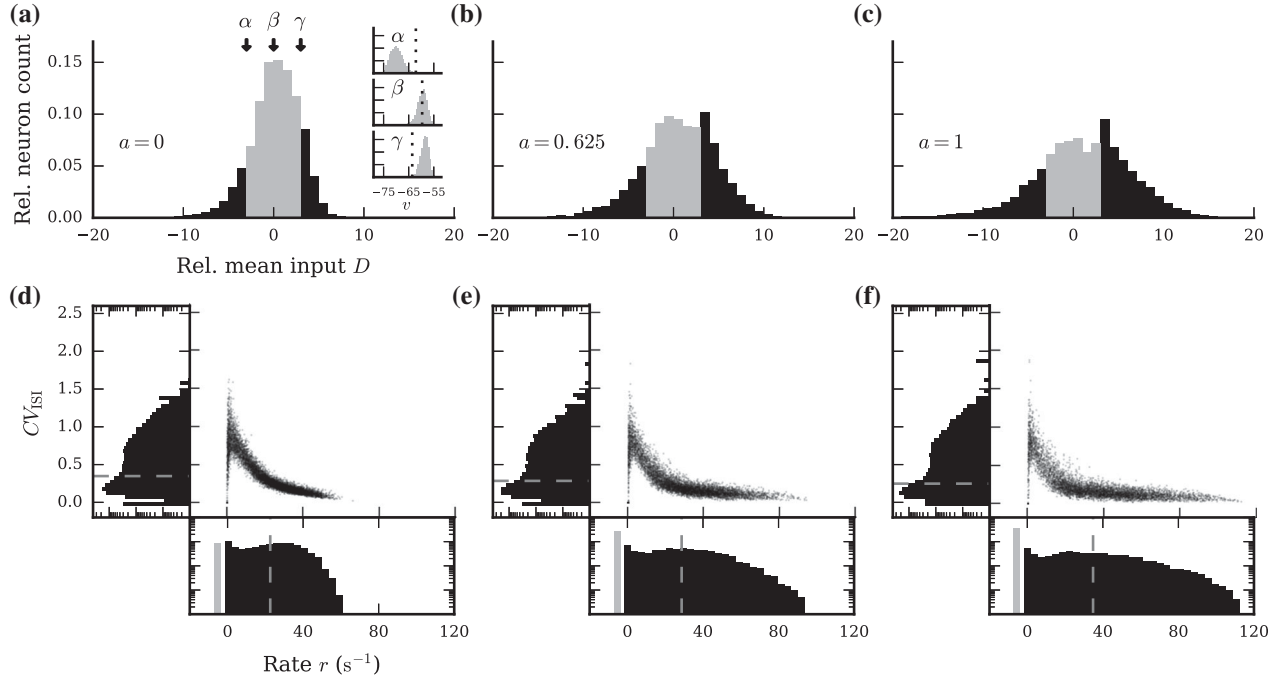


FIG. 6. Modulation of network heterogeneity by leak-conductance calibration (see Sec. II E). Input (top row) and firing statistics (bottom row) in the intact recurrent networks (FB scenarios) for fully calibrated [(a),(d) $a = 0$], partially calibrated [(b),(e) $a = 0.625$], and uncalibrated neurons [(c),(f) $a = 1$]. (a)–(c) Effect of calibration on input statistics. Distributions of relative mean input $D = (\bar{v} - \Theta)/\sigma(v)$ [distance of time-averaged free membrane potential \bar{v} from firing threshold Θ in units of the standard deviation $\sigma(v)$] across the population of neurons. Gray areas in (a)–(c) highlight $[-3, 3]$ intervals, containing 74%, 53%, and 42% of the total mass of the distribution, respectively. Inset in (a): Distributions of free membrane potentials v for three neurons α , β , and γ , with $D = -3$, $D = 0$, and $D = 3$ [arrows in (a)], respectively. Dotted lines mark threshold potentials that may vary due to fixed-pattern noise. (d)–(f) Effect of calibration on spike-train statistics. Joint (scatter plots) and marginal distributions of single-neuron firing rates r (horizontal histograms, log-linear scale) and coefficients of variation CV_{ISI} of interspike intervals (vertical histograms, log-linear scale). Dashed lines mark mean of firing rate (22.6, 28.7, 34.8 s^{-1}) and CV_{ISI} distributions (0.35, 0.28, 0.25), respectively. Gray bars (bottom panels) represent fractions of silent neurons. Data obtained from $M = 50$ different network realizations. Percentage of dead neurons: 8%, 26%, 37%.

positive contribution from shared-input correlations. Average input correlations are, therefore, significantly reduced [black symbols in Fig. 7(a)]. As both shared-input and spike-train correlations scale with the inverse of the network size [N^{-1} ; light gray curve in Fig. 7(a) and inset in Fig. 7(b), respectively] [94], this suppression of correlations in the FB (and FB_{replay}) case is observed for all investigated network sizes N . Note that output correlations are negative even though input correlations are positive. This effect is predicted by theory and also observed in linear network models as well as LIF-network simulations on conventional computers (see Fig. 9, Sec. IV, and Supplemental Material Secs. IV and V [78]).

B. Effect of heterogeneity on decorrelation

In neural networks implemented in analog neuromorphic hardware, neuron and synapse parameters vary significantly across the population of cells (fixed-pattern noise; see Sec. II B). For a population of mutually unconnected neurons with distributed parameters, injection of a constant

(suprathreshold) input current leads to a distribution of response firing rates (Fig. 4). In this study, we consider the width of this firing-rate distribution as a representation of neuron heterogeneity. It is systematically varied by calibration of leak conductances. The extent of heterogeneity is quantified by the calibration parameter a ($a = 1$ and $a = 0$ correspond to the uncalibrated and the fully calibrated system, respectively; for details, see Sec. II E). For an unconnected population of neurons subject to constant input, the width of the firing-rate distribution increases monotonically with a .

As shown in Fig. 6, the level of heterogeneity (i.e., the calibration state a) is clearly reflected in the activity of the recurrent network (FB case). Both the width of the distribution of mean free membrane potentials [Figs. 6(a)–6(c)] as well as the width of the firing-rate distribution increase with a [Figs. 6(d)–6(f), horizontal histograms]. In the uncalibrated system ($a = 1$), a substantial fraction of neurons is predominantly driven by constant suprathreshold input currents and therefore generates highly regular spike trains ($CV_{ISI} \approx 0$) with high firing rates ($r > 60 s^{-1}$). Simultaneously, about

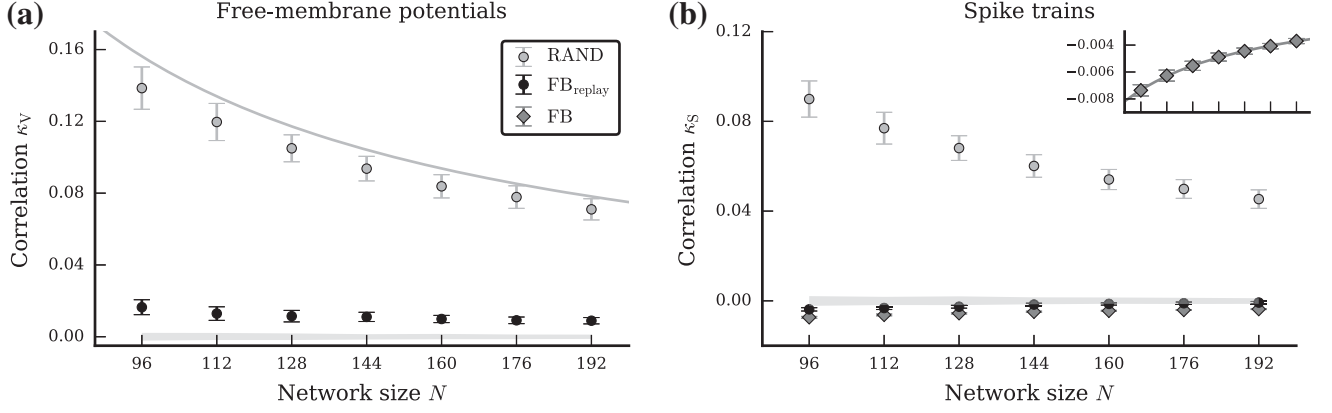


FIG. 7. Dependence of population-averaged input correlations (a) and spike-train correlations (b) on the network size N , for the intact network (FB, dark gray diamonds), the FB_{replay} (black circles), and the RAND (light gray circles) cases (fixed in-degree $K = 15$). Symbols and error bars denote mean and 1 standard deviation, respectively, across $M = 100$ network realizations (error bars are partly covered by markers). The gray curve in (a) depicts shared-input correlations in a homogeneous network [Eq. (2)]. The inset in (b) shows a magnified view of the spike-train correlations in the FB case (dark gray diamonds) with a power-law fit $\sim N^{-1}$ (dark gray curve). The light gray horizontal band represents mean ± 3 standard deviations of (spurious) correlations in surrogate data where correlations are removed. Note that free membrane potentials cannot be recorded in the FB case (see Sec. II). Hence, there are no gray diamonds in (a).

37% of the neurons are silent ($r = 0 \text{ s}^{-1}$). Neurons with intermediate firing rates ($0 < r < 15 \text{ s}^{-1}$), however, show quite irregular activity ($CV_{\text{ISI}} > 0.5$). After calibration, the firing-rate distribution is narrowed. For $a = 0$, the fraction of silent neurons is reduced to about 8%. Maximum rates are limited to $< 66 \text{ s}^{-1}$. Note that our calibration routine compensates only for the distribution of neuron parameters, but not for the heterogeneity in synapse properties (synaptic weights, synaptic time constants; see Sec. IV). Even for the fully calibrated network ($a = 0$), the firing-rate distribution

is therefore still broad. In the RAND case, we obtain similar firing-rate and ISI statistics as in the FB case (see Supplemental Material, Sec. VI [78]).

For all levels of heterogeneity attainable by our calibration procedure ($a \in [0, 1]$), input and output correlations are significantly suppressed by the recurrent-network dynamics (cf. black and dark gray versus light gray symbols in Fig. 8). In a homogeneous, random (Erdős-Rényi) network with fixed in-degree K and linear subthreshold dynamics, the contribution of

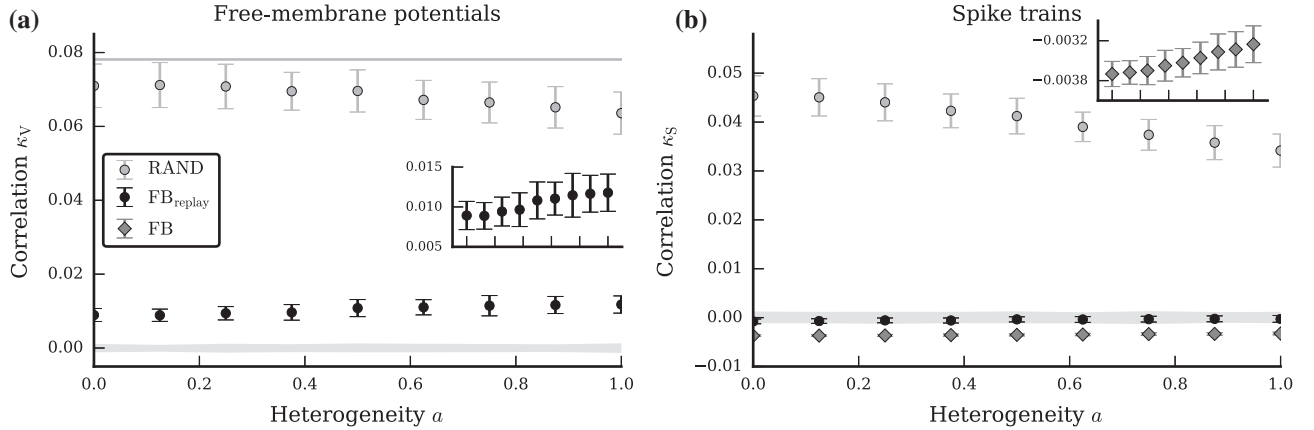


FIG. 8. Dependence of population-averaged input correlations (a) and spike-train correlations (b) on the heterogeneity of the neuromorphic substrate for the intact network (FB, dark gray diamonds), the FB_{replay} (black circles), and the RAND (light gray circles) cases. Symbols and error bars denote mean and 1 standard deviation, respectively, across $M = 100$ network realizations (error bars are partly covered by markers). The gray line in (a) depicts shared-input correlations in a homogeneous network [Eq. (2)]. Insets in (a) and (b) depict magnified views of input correlations in the FB_{replay} case and spike-train correlations in the FB case, respectively. The light gray horizontal band represents mean ± 3 standard deviations of (spurious) correlations in surrogate data where correlations were removed. Note that free membrane potentials cannot be recorded in the FB case (see Sec. II). Hence, there are no gray diamonds in (a).

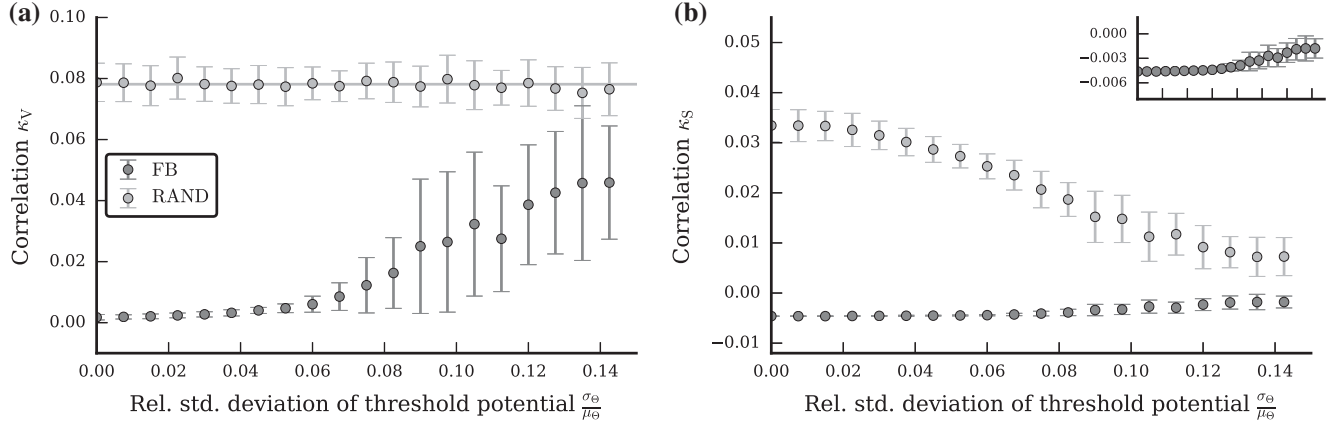


FIG. 9. Dependence of population-averaged input correlations (a) and spike-train correlations (b) on the width of threshold distributions in networks simulated with NEST [101] and PyNN [88], for the intact network (FB, dark gray circles) and the RAND (light gray circles) case. Symbols and error bars denote mean and standard deviation, respectively, across $M = 30$ network realizations (error bars are partly covered by markers). The gray line in (a) depicts shared-input correlations in a homogeneous network [Eq. (2)]. The inset in (b) shows a magnified view of the spike-train correlations in the FB case. Note that in simulations the $\text{FB}_{\text{replay}}$ is identical to the FB case, and is hence not shown. For details, see Supplemental Material, Sec. IV [78].

shared input to the input (free-membrane-potential) correlation is given by the network connectivity K/N [Eq. (2)] [5] [thin light gray curves in Figs. 7(a) and 8(a)]. Nonlinearities in synaptic and/or spike-generation dynamics [94] as well as heterogeneity in neuron (and synapse) parameters lead to a suppression of this contribution [Eq. (3)] [69]. Here, we refer to this type of decorrelation as feedforward decorrelation. In fact, in our setup the input and spike-train correlations in the RAND case decrease with increasing heterogeneity (light gray symbols in Fig. 8). Even in the fully calibrated case input correlations are slightly smaller than K/N [gray symbols versus thin gray curve in Fig. 8(a)] due to remaining heterogeneities. Spike-train correlations decrease slightly faster with increasing heterogeneity than input correlations. These observations indicate that both the synaptic integration and spike generation are affected by heterogeneities on hardware. To illustrate the different effects of heterogeneity in synaptic integration and spike generation, we perform network simulations on conventional computers where we distribute either firing thresholds (Fig. 9) or synaptic weights (see Supplemental Material, Fig. 4 [78]). In the RAND case, the decrease of input correlations on the hardware can be attributed to an increase in heterogeneity in parameters affecting synaptic integration [compare light gray symbols in Fig. 8(a) and Supplemental Material Fig. 4(a) [78]]. Heterogeneity in spike thresholds, in contrast, does not affect input correlations in the RAND scenario, but strongly reduces spike-train correlations (light gray symbols in Fig. 9). Overall, feedforward decorrelation, i.e., the suppression of correlations in the RAND case, becomes more effective in networks with heterogeneous cell parameters.

Despite this enhancement of feedforward decorrelation, input and output correlations increase with the level of heterogeneity in the presence of an intact feedback loop (FB and $\text{FB}_{\text{replay}}$ scenarios, black and dark gray symbols in Fig. 8). We attribute this effect to a weakening of the effective feedback loop in the recurrent circuit: In heterogeneous networks with broad firing-rate distributions, neurons firing with low or high rates, corresponding to mean inputs far below or far above firing threshold [see Figs. 6(a)–6(c)], are less sensitive to input fluctuations than moderately active neurons (see Supplemental Material Fig. 2 [78]). Hence, they contribute less to the overall feedback. In consequence, feedback decorrelation is impaired by heterogeneity (see also Sec. IV).

IV. DISCUSSION

We show that inhibitory feedback effectively suppresses correlations in heterogeneous recurrent neural networks of leaky integrate-and-fire neurons with nonlinear subthreshold dynamics, emulated on analog neuromorphic hardware (Spikey) [18,77]. Both input and output correlations are substantially smaller in networks with intact feedback loop (FB), as compared to the case where the feedback is replaced by randomized input while preserving the connectivity structure and presynaptic firing rates (RAND). Our results, hence, show that active decorrelation of network activity by inhibitory feedback [4,5] is a general phenomenon that can be observed in realistic, highly heterogeneous networks with nonlinear interaction and sufficiently strong negative feedback. Moreover, the study serves as a proof of principle that network activity can be efficiently decorrelated even on heterogeneous hardware, which can be exploited in

functional applications, e.g., in the neuromorphic algorithms developed by Pfeil *et al.* [18] and Schmuker *et al.* [19].

Functional neural architectures often rely on stochastic dynamics of its constituents or on some form of background noise (see, e.g., Refs. [18,19,102]). Deterministic recurrent neural networks with inhibitory feedback could provide decorrelated noise to such functional networks, both in artificial as well as in biological substrates. In neuromorphic-hardware applications, these “noise networks” could thereby replace conventional random-number generators and avoid a costly transmission of background noise from a host computer to the hardware substrate (which may be particularly relevant for mobile applications with low power consumption; see Supplemental Material, Sec. I [78]). It needs to be investigated, however, how well functional stochastic circuits perform in the presence of such network-generated noise.

Partial calibration of hardware neurons allows us to modulate the level of network heterogeneity and, therefore, to systematically study its effect on correlations in the network activity. The analysis reveals two counteracting contributions: As shown in previous studies (e.g., Ref. [69]), neuron heterogeneity decorrelates (shared) feedforward input (feedforward decorrelation). On the other hand, however, heterogeneity impairs feedback decorrelation (see next paragraph). In our network model, this weakening of feedback decorrelation is the dominating factor. Overall, we observe a slight increase in correlations with increasing level of heterogeneity. We cannot exclude that feedforward decorrelation may play a more significant role for different network configurations (e.g., different connection strengths or network topologies, different structure of external inputs, different types of heterogeneity). Our study demonstrates, however, that heterogeneity is not necessarily suppressing correlations in recurrent systems. In this context, it would be interesting to investigate the interplay of signal and noise correlations in the presence of network heterogeneities in recurrent systems. We leave this intriguing topic to future studies.

As shown in Ref. [5], feedback decorrelation in recurrent networks becomes more (less) efficient with increasing (decreasing) strength of the effective negative feedback. For networks of spiking neurons, the effective connection strength w_{ij} between two neurons j and i corresponds to the total number of extra spikes emitted by neuron i in response to an additional input spike generated by neuron j (see, e.g., Ref. [103]). Assuming that the effect of a single additional input spike is small, the effective connectivity can be obtained by linear-response theory [94]. Note that the effective weights w_{ij} depend on the working point, i.e., the average firing rates of all pre- and postsynaptic neurons [mathematically, w_{ij} is given by the derivative of the stationary response firing rate $r_i = \phi_i(r_1, \dots, r_j, \dots, r_N)$ of neuron i with respect to

the input firing rate r_j , evaluated at the working point; for details, see Ref. [5]]. Neurons firing at very low or very high rates are typically less sensitive to input fluctuations than neurons firing at intermediate rates [due to the shape of the response function $\phi_i(r_1, \dots, r_N)$]. Their dynamical range is reduced. In consequence, they hardly mediate feedback in a recurrent network. In heterogeneous networks with broad distributions of firing rates, the number of these insensitive neurons is increased. Hence, the effective feedback is weakened (see Supplemental Material, Sec. III [78]). We can qualitatively reproduce this effect of heterogeneity on correlations in recurrent networks (FB case) by means of a simplified linear rate model where increasing heterogeneity is described as a decrease in the effective-weight amplitudes (see Supplemental Material, Sec. V [78]). A more quantitative analysis requires an explicit mapping of the synaptic weights in the LIF-neuron network to the effective weights of the linear model (as in, e.g., Ref. [104]) in the presence of distributed firing rates. We commit this task to future studies. Note that the rate dependence of the effective weights and the resulting effects on correlations are consistent with our observation that neuron pairs with very low firing rates exhibit spike-train correlations close to zero, whereas pairs with high firing rates are positively correlated (see Supplemental Material, Sec. VII [78]). Pairs with one neuron firing at an intermediate rate often exhibit negative spike-train correlations. As shown in Refs. [4,5], these negative spike-train correlations are essential for compensating the positive contribution of shared inputs to the total input correlation (at least in purely inhibitory networks). Narrowing the firing rate distribution (e.g., by calibration of hardware neurons) increases the number of neurons contributing to the negative feedback, which, in turn, leads to more neuron pairs with negative spike-train correlations and, therefore, to smaller overall correlations.

Seemingly contrary to our findings, Bernacchia and Wang [76] report a decrease in correlations with increasing level of heterogeneity. The results of their study are obtained for a linear network model, which can be considered the outcome of the linearization procedure described above. Hence, the connectivity of their model corresponds to an effective connectivity (see above). Their study neglects the rate (working-point) dependence of the effective weights and, therefore, cannot account for the effect of firing-rate heterogeneity. In Ref. [76], heterogeneity is quantified by the variance of the (effective) weight matrix [Eqs. (2.2) and (2.4) in Ref. [76]]. For sparse connectivity matrices (with a large number of zero elements), the variance of the weight matrix reflects not only the width of the nonzero-weight distribution, but also its mean [Eq. (2.4) in Ref. [76]]. For networks of nonlinear spiking neurons, heterogeneities in neuron and/or synapse parameters broaden the distribution of nonzero effective

weights, but may simultaneously reduce its mean (see above, Supplemental Material Sec. III [78], and Refs. [48,104]). Hence, the variance of the full weight matrix may decrease (for illustration, see Supplemental Material, Fig. 9 [78]). In other words, increasing heterogeneity in the nonlinear system may correspond to decreasing heterogeneity in the linearized system. A direct test of this hypothesis requires an explicit linearization of the nonlinear heterogeneous system.

The results of this study are obtained by network emulations on analog neuromorphic hardware. We reproduce the main findings by means of conventional computer simulations of LIF-neuron networks with distributed firing thresholds (see Fig. 9). The focus on threshold heterogeneity allows us to isolate the effect of firing-rate distributions on correlations. It does not affect shared-input correlations (see Sec. II F). Although networks simulated on conventional computers and those emulated on the neuromorphic hardware differ in several respects (e.g., in the exact implementation of heterogeneity or the synapse model; compare Tables I and II in Supplemental Material [78] to Tables I and II, respectively), the qualitative results are very similar: In networks with intact feedback loop, input and output correlations are substantially reduced (as compared to the case where the feedback is replaced by randomized input), but increase with the extent of heterogeneity. As predicted by the theory for homogeneous inhibitory networks, we observe positive input correlations and negative output correlations [see Eq. (21) in Ref. [5] and in the paragraph that follows; see also Ref. [105] and Supplemental Material Sec. V [78]]. Further, note that heterogeneity in neuron parameters does not “average out” in larger networks. Upscaling the network size by a factor of 25 ($N = 4800$, in-degree $K = 375$) yields smaller spike-train correlations, but the qualitative results are similar to those obtained for the smaller network ($N = 192$, $K = 15$) emulated on the Spikey chip (compare Fig. 8 to Supplemental Material Fig. 3 [78]).

In networks with intact feedback loop (FB and FB_{replay} scenarios), the precise spatiotemporal structure of spike trains arranges such that the self-consistent input and output correlations are suppressed. Perturbations of this structure in the local input typically lead to an increase in correlations [5]. In this study, we demonstrate this by replaying spiking activity after randomization of spike times, i.e., by replacing the time of each input spike by a random number uniformly drawn from the full emulation time interval $[0, T)$ (RAND case). However, even subtle modifications of input spike trains, such as random dither of spike times by a few milliseconds, lead to an increase of correlations. On the neuromorphic hardware, replay of spike trains is not entirely reproducible (see Sec. II D and Supplemental Material, Sec. IX [78]). Hence, spike-train correlations measured in the FB_{replay} mode are slightly larger than in the

FB case. We would expect the same effect on the input side (free membrane potentials). Because of hardware limitations, however, we can measure input correlations only in replay mode (FB_{replay} or RAND), but not in the fully connected network (FB). Therefore, all reported input correlations are likely to be slightly overestimated. In conventional network simulations, we mimic the effect of unreliable replay by input-spike dithering and, indeed, find a gradual increase in input and output correlations (see Supplemental Material, Fig. 6 [78]). These results seem to be contrary to the study by Rosenbaum and Josic [106], in which synaptic noise leads to a decrease of output correlations in a feedforward scenario. In our case, spike-train correlations, which suppress shared-input correlations, are removed by dithering spikes, thereby increasing correlations on the output side. In Ref. [106], in contrast, spike-train correlations are always zero, and shared-input correlations are decreased by synaptic failure, explaining the decreased output correlations. We attribute this contradiction to the missing feedback loop in their system, and expect correlations to increase in recurrent networks subject to similar perturbations.

Despite the imperfect replay of input spikes, the decorrelation effect is clearly visible in hardware emulations, both on the input and on the output side. The reproducibility of emulations on neuromorphic hardware could be improved by stabilizing the environment of the system, e.g., the chip temperature or the support electronics (under development). Analog hardware, however, will never reach the level of reproducibility of digital computers. But note that, similar to analog hardware, biological neurons exhibit a considerable amount of trial-to-trial variability, even under controlled *in vitro* conditions [93]. Thus far, the details of how neuronal noise, for example, stochastic synapses (spontaneous postsynaptic events, stochastic spike transmission, synaptic failure [107]), affects correlations in recurrent neural circuits remain unclear.

Although different Spikey chips exhibit different realizations of fixed-pattern noise, they show a comparable extent of heterogeneity and yield results that are qualitatively similar to those presented in this article (see Supplemental Material, Sec. VIII [78]).

We show that negative feedback in recurrent circuits can efficiently suppress correlations, even in highly heterogeneous systems such as the analog neuromorphic architecture Spikey. Correlations can be further reduced by minimizing the level of network heterogeneity. In this study, we reduce the level of heterogeneity through calibration of neuron parameters in the unconnected case (see Sec. II E). The calibration could, in principle, be improved by calibrating neuron (and possibly synapse) parameters in the full recurrent network. Such calibration procedures are, however, time-consuming and cumbersome. In biological substrates, homeostasis mechanisms [58,108] keep neurons in a responsive

regime and reduce the level of firing-rate heterogeneity in a self-regulating manner. Future neuromorphic devices could mimic this behavior, thereby reducing the necessity of time-consuming calibration procedures. Alternatively, the analog circuits could be optimized to reduce fixed-pattern noise. This would likely require the allocation of more chip resources, hence reducing the network size per chip area.

For simplicity, this work focuses on purely inhibitory networks (as in Ref. [11]). This demonstrates that decorrelation by inhibitory feedback does not rely on a dynamical balance between excitation and inhibition (note that the external “excitatory” drive is constant in our model) [5,6]. Previous studies have shown that, for the homogeneous case, decorrelation by inhibitory feedback is a general phenomenon, which also occurs in excitatory-inhibitory networks, provided the overall inhibition is sufficiently strong (which is typically the case to ensure stability) [4–6,76]. For the heterogeneous case, computer simulations of excitatory-inhibitory networks show qualitatively the same results as purely inhibitory networks (compare Fig. 8 to Supplemental Material Fig. 5 [78]), confirming that our results generalize to the case of mixed excitatory-inhibitory coupling.

Similar to our study, Giulioni *et al.* [109] use a theory-guided approach to implement, verify, and investigate network dynamics on analog neuromorphic hardware. In their study, an attractor network is implemented that is inspired by a mean-field model. Because of heterogeneities in synaptic efficacies on the hardware, stability analysis of attractor states requires the authors to measure effective response functions of populations of hardware neurons. To this end, they replace recurrent connections in one population of neurons by external input. This allows them to measure the firing rate of the population as a function of the external input, while the activity of the population is in equilibrium with that of other recurrently connected populations. This study represents another example illustrating that investigations of actual hardware emulations are a prerequisite for successful application of analog neuromorphic hardware.

This study demonstrates that the Spikey system has matured to a level that permits its use as a tool for neuroscientific research. For the results we present in this study, we record in total 10^{11} membrane-potential and spike-train samples, representing more than 100 days of biological time. Because of the 10^4 -fold acceleration of the Spikey chip, this corresponds to less than 15 min in the hardware-time domain. Interfacing the hardware system, however, reduces the acceleration to an approximately 50-fold speed-up (Fig. 10). The translation between the network description and its hardware representation claims the majority of execution time, more than the network emulation and the transfer of data to and from the hardware system together. Encoding and decoding spike

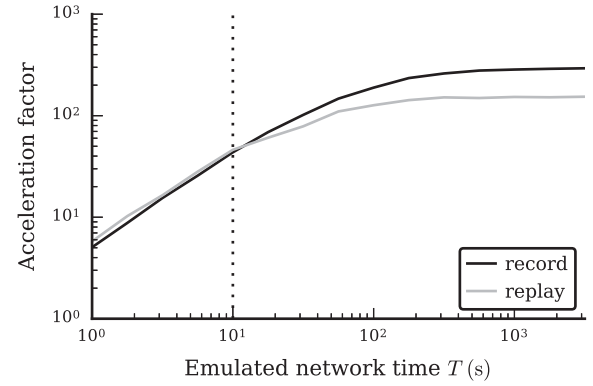


FIG. 10. Acceleration factor as a function of emulated network time T for the record (black line) and the replay cases (gray line). The acceleration factor is defined as the ratio between the emulated network time T (in biological time) and the execution time (wall clock time). In the record case, a network realization is generated on the host computer and uploaded to the chip. During the subsequent emulation, spike trains are recorded. In the replay case, spikes are replayed and the membrane potential of one neuron is recorded with full sampling frequency (9.6 kHz). The execution time covers the full data flow from a network description in PyNN to the emulation on the Spikey system and back to the network representation in PyNN. The time-averaged population firing rate is $\bar{r} = (23.7 \pm 0.2) \text{ s}^{-1}$. The vertical dashed line depicts the run time used in this study. The hardware system has to be initialized once before usage ($< 1 \text{ s}$), which is not considered here.

times on the host computer is particularly expensive. Obviously, the system could be optimized by processing the data directly on the hardware, or by choosing a data representation that is closer to the format used on the Spikey chip, but this would impair user friendliness, and hence, the effectiveness of prototyping. While the Spikey system permits the monitoring of the spiking activity of all neurons simultaneously, access to the membrane potentials is limited to a single (albeit arbitrary) neuron in each emulation run. Monitoring of membrane potentials of a population of n neurons therefore requires n repetitions of the same emulation. Extending the hardware system to enable access to the membrane potentials of at least two neurons simultaneously would allow for a direct observation of input correlations in the intact network (and thereby avoid problems with replay reproducibility; see above) and reduce execution time (the Spikey chip itself permits recording of up to eight neurons in parallel, the support electronics, however, does not). While the Spikey system does not significantly outperform conventional computers in terms of computational power, emulations on this system are much more energy efficient (Supplemental Material, Sec. I [78]). A substantial increase of computational power is expected for large systems exploiting the scalability of this technology without slow-down [110].

ACKNOWLEDGMENTS

We would like to thank Matthias Hock, Andreas Hartel, Eric Müller, and Christoph Koke for their support in developing and building the Spikey system, as well as Mihai Petrovici and Sebastian Schmitt for fruitful discussions. This research is partially supported by the Helmholtz Association portfolio theme SMHB, the Jülich Aachen Research Alliance (JARA), EU Grant No. 269921 (BrainScaleS), and EU Grant No. 604102 (Human Brain Project, HBP). We acknowledge financial support by Deutsche Forschungsgemeinschaft and Ruprecht-Karls-Universität Heidelberg within the funding programme Open Access Publishing. All network simulations carried out with NEST (<http://www.nest-simulator.org>).

T.P. and J.J. contributed equally to this work.

APPENDIX A: NETWORK DESCRIPTION

Details on the network model as well as parameter values are provided in Tables I and II, respectively.

TABLE I. Description of the network model (according to Ref. [111]).

| A | | Model summary | |
|-----------------------|--|--|--|
| Populations | One (inhibitory) | | |
| Topology | None | | |
| Connectivity | Random convergent connections (fixed in-degree) | | |
| Neuron model | Leaky integrate-and-fire (LIF), fixed firing threshold, fixed absolute refractory time | | |
| Channel models | None | | |
| Synapse model | Exponentially decaying conductances, fixed delays | | |
| Plasticity | None | | |
| External input | Resting potential higher than threshold (= constant current) ($E_1 > \Theta$) | | |
| Measurements | Spikes and membrane potentials from all neurons | | |
| Other | No autapses, no multapses | | |
| | | | |
| B | | Populations | |
| Name | Elements | Size | |
| I | LIF neuron | N | |
| | | | |
| C | | Connectivity | |
| Source | Target | Pattern | |
| I | I | Random convergent connections, in-degree K | |
| | | | |
| D | | Neuron and synapse model | |
| Type | Leaky integrate-and-fire, exponentially decaying conductances | | |
| Subthreshold dynamics | Subthreshold dynamics $[t \in (t^*, t^* + \tau_{\text{ref}})]$: $C_m(d/dt)v(t) = -g_l[v(t) - E_l] - g_{\text{syn}}(t)[v(t) - E_{\text{inh}}]$ Reset and refractoriness $[t \in (t^*, t^* + \tau_{\text{ref}})]$: $v(t) = v_{\text{reset}}$. This model is emulated by analog circuitry on the Spikey chip [13]. | | |
| Conductance dynamics | For each presynaptic spike at time t^* ($t > t^* + d$): $g_{\text{syn}}(t) \approx J \exp[-(t - t^* - d)/(\tau_{\text{syn}})]\Theta(t)$, with $J = w_{\text{hw}}g_{\text{max}}$ and Heaviside function $\Theta(t)$. This model is emulated by analog circuitry on the Spikey chip [18]. | | |
| Spiking | If $v(t^*-) < \Theta \wedge v(t^*+) \geq \Theta$: emit spike with time stamp t^* | | |

TABLE II. Parameter values for the network model described in Table I. Bold numbers indicate default values. Italic numbers indicate target values not considering fixed-pattern noise. Leak conductances g_l are adjusted in the calibration process (see Sec. II E). τ_m^{eff} , $\tau_{\text{syn}}^{\text{eff}}$, d^{eff} , and $V_{\text{max}}^{\text{eff}}$ describe effective values measured from spike-triggered averages in RAND emulations as described in Supplemental Material, Sec. X. Effective values denote the median across synapses and trials followed by the 25th and 75th percentiles in brackets.

| B | | | <i>Populations</i> |
|----------------------------------|--|---|---------------------|
| Name | Values | Description | |
| N | {96, 112, 128, 144, 160, 176, 192 } | Network size | |
| C | | | <i>Connectivity</i> |
| Name | Values | Description | |
| K | 15 | Number of presynaptic partners | |
| D | | | <i>Neuron</i> |
| Name | Values | Description | |
| v_{reset} | −80 mV | Reset potential | |
| E_l | −52 mV | Resting potential | |
| Θ | −62 mV | Firing threshold | |
| E_{inh} | −80 mV | Inhibitory reversal potential | |
| τ_{ref} | 1 ms | Refractory period | |
| C_m | 0.2 nF | Membrane capacitance | |
| g_l | Tuned during calibration process | Leak conductance | |
| τ_m^{eff} | Uncalibrated: 10.50(5.94,17.90) ms Calibrated: 10.77(6.22,18.30) ms | Effective membrane time constant | |
| D | | | <i>Synapse</i> |
| Name | Values | Description | |
| g_{max} | In the order of 1 nS | Conductance amplitude | |
| w_{hw} | 3 | Synaptic weight (in hardware values $\in [0, 15]$) | |
| $V_{\text{max}}^{\text{eff}}$ | Uncalibrated: −5.57(−12.50, −2.62) mV Calibrated: −6.05(−14.37, −2.92) mV | Effective post-synaptic potential prefactor | |
| $\tau_{\text{syn}}^{\text{eff}}$ | Uncalibrated: 3.78(2.52,5.58) ms Calibrated: 4.17(2.76,6.18) ms | Effective synaptic-current time constant | |
| d^{eff} | Uncalibrated: 2.16(1.91,2.48) ms Calibrated: 2.26(1.93,2.55) | Effective synaptic delay | |
| Other | | | <i>Software</i> |
| Name | Values | Description | |
| SpikeyHAL | 9e86d11c | git revision | |
| PyNN | 2fe40b43 | git revision | |
| vmodule | 76ef3b44 | git revision | |
| logger | 826c5ed6 | git revision | |
| Other | | | <i>Hardware</i> |
| Name | Values | Description | |
| | Chip 508 (version 5) | Chip used for main manuscript | |
| | Chip 503, 504 (version 5) and 603, 605, 666 (version 4) | Chips used for Supplemental Material [78] | |

APPENDIX B: DESCRIPTION OF DATA ANALYSIS

A summary of the quantities and the data analysis used in this study is provided in Table III. Parameter values for the data analysis are given in Table IV.

TABLE III. Summary of the data analysis. Here, $i \in [1, N]$, $X \in S, V$.

| A | Analysis measures |
|--|--|
| Measure | Details |
| Spike density | $\xi_i(t) = \sum_k \delta(t - t_i^k)$ |
| Spike train | $s_i(t_k) = \text{number of spikes of neuron } i \text{ per bin } [k\Delta t, (k+1)\Delta t]$ |
| Population activity | $\bar{s}(t) = (1/N) \sum_i s_i(t)$ |
| Time-averaged population activity | $\bar{r} = \langle \bar{s}(t) \rangle_t$ |
| Membrane potential | $v_i(t_k) = \text{membrane potential of neuron } i \text{ in bin } [k\Delta t_m, (k+1)\Delta t_m]$ |
| (Finite time) Fourier transform | $X_i(f) = \mathfrak{F}[x_i(t)](f) = \int_0^T dt x_i(t) e^{-2\pi i f t}$ (with inverse \mathfrak{F}^{-1}) |
| (Single unit) power spectrum | $A_i(f) = (1/T) X_i^*(f) X_i(f)$ |
| Population-averaged power spectrum | $A(f) = (1/N) \sum_i A_i(f)$ |
| Population power spectrum | $\bar{A}(f) = (1/T) [\sum_i S_i^*(f)] [\sum_j S_j(f)]$ |
| Pairwise cross spectrum | $C_{ij} = (1/T) X_i^*(f) X_j(f), i \neq j$ |
| Population-averaged cross spectrum | $C(f) = \{1/[N(N-1)]\} \sum_{i \neq j} C_{ij}(f) \equiv \{1/[N(N-1)]\} [\bar{A}(f) - NA(f)]$ [note: $C(f) \in \mathbb{R}$] |
| Sliding window filter | $X(f) \rightarrow X(f) * H(f),$ with $H(f) = [1/(f_1 - f_0)] \Theta(f - f_0) \Theta(f_1 - f)$ |
| Coherence | $\kappa(f) = [C(f)/A(f)]$ |
| Low-frequency coherence | $\kappa_X = [1/(f_{\max} - f_{\min})] \int_{f_{\min}}^{f_{\max}} df \kappa(f)$ |
| Population-averaged cross-correlation function | $c(\tau) = \{1/[N(N-1)]\} \sum_{i \neq j} \langle s_i(t) s_j(t + \tau) \rangle_t \equiv \mathfrak{F}^{-1}[C(f)](\tau)$ |
| Time average | $\langle \dots \rangle_t$ |

TABLE IV. Summary of analysis parameters (default values in bold).

| A | Analysis parameters | |
|----------------------------------|---|--|
| Parameter | Description | Values |
| Δt | Bin size for spike trains | 1 ms |
| Δt_m | Bin size for membrane-potential traces | 0.52 ms |
| T_{warmup} | Initial warm-up time (not considered in analysis) | 1 s |
| T | Network emulation time | 10 s |
| M | Number of network realizations | {50, 100 } |
| ΔF | Width of sliding window | 1 Hz |
| $f_{\text{min}}, f_{\text{max}}$ | Interval boundaries for low-frequency coherence | 0.1 Hz, 20 Hz |
| a | Calibration state | $\{0, \frac{1}{8}, \frac{2}{8}, \frac{3}{8}, \frac{4}{8}, \frac{5}{8}, \frac{6}{8}, \frac{7}{8}, \mathbf{1}\}$ |

- [1] A. Gierer and H. Meinhardt, *Biological Pattern Formation Involving Lateral Inhibition*, Lect. Math. Life Sci. **7**, 163 (1974).
- [2] M. L. Rosenzweig, *Competitive Speciation*, *Biological Journal of the Linnean Society* **10**, 275 (1978).
- [3] L. Neel, *Antiferromagnetism and Ferrimagnetism*, *Proc. Phys. Soc. London Sect. A* **65**, 869 (1952).
- [4] A. Renart, J. De La Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris, *The Asynchronous State in Cortical Circuits*, *Science* **327**, 587 (2010).
- [5] T. Tetzlaff, M. Helias, G. Einevoll, and M. Diesmann, *Decorrelation of Neural-Network Activity by Inhibitory Feedback*, *PLoS Comput. Biol.* **8**, e1002596 (2012).
- [6] M. Helias, T. Tetzlaff, and M. Diesmann, *The Correlation Structure of Local Cortical Networks Intrinsically Results from Recurrent Dynamics*, *PLoS Comput. Biol.* **10**, e1003428 (2014).
- [7] K. Pearson, *On Lines and Planes of Closest Fit to Systems of Points in Space*, *Philos. Mag. Ser. 6* **2**, 559 (1901).
- [8] G. Ginis and C.-N. Peng, *Alien Crosstalk Cancellation for Multipair Digital Subscriber Line Systems*, *EURASIP J. Adv. Signal Process.* **1**, 1 (2006).
- [9] J. Benesty, T. Gänslar, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*, 1st ed. (Springer, Berlin, 2001).
- [10] E. Trichina, M. Bucci, D. De Seta, and R. Luzzi, *Supplemental Cryptographic Hardware for Smart Cards*, *IEEE Micro* **21**, 26 (2001).
- [11] D. Mar, C. Chow, W. Gerstner, R. Adams, and J. Collins, *Noise Shaping in Populations of Coupled Model Neurons*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10450 (1999).
- [12] C. Mead, *Neuromorphic Electronic Systems*, *Proc. IEEE* **78**, 1629 (1990).
- [13] G. Indiveri *et al.*, *Neuromorphic Silicon Neuron Circuits*, *Front. Neurosci.* **5**, 1 (2011).
- [14] *SenseMaker: A Multi-Sensory, Task-Specific, Adaptable Perception System*, http://cordis.europa.eu/project/rcn/62746_en.html.
- [15] *FACETS: Fast Analog Computing with Emergent Transient States*, <http://www.facets-project.org>.
- [16] *BrainScaleS*, <http://www.brainscales.eu>.
- [17] *Human Brain Project*, <http://www.humanbrainproject.eu>.
- [18] T. Pfeil, A. Grübl, S. Jeltsch, E. Müller, P. Müller, M. A. Petrovici, M. Schmuker, D. Brüderle, J. Schemmel, and K. Meier, *Six Networks on a Universal Neuromorphic Computing Substrate*, *Front. Neurosci.* **7**, 1 (2013).
- [19] M. Schmuker, T. Pfeil, and M. P. Nawrot, *A Neuromorphic Network for Generic Multivariate Data Classification*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2081 (2014).
- [20] J. Marienborg, T. Lande, and M. Hovén, in *Proceedings of the IEEE International Symposium on Circuits and Systems, 2002 (ISCAS 2002)* (IEEE, New York, 2002), Vol. 5, pp. V-73–V-76.
- [21] A. Kikombo, T. Asai, T. Oya, A. Schmid, Y. Leblebici, and Y. Amemiya, in *Proceedings of the International Joint Conference on Neural Networks, 2009 (IJCNN 2009)* (IEEE Computer Society, Atlanta, GA, 2009), pp. 1600–1605.
- [22] A. S. Ecker, P. Berens, A. S. Tolias, and M. Bethge, *The Effect of Noise Correlations in Populations of Diversely Tuned Neurons*, *J. Neurosci.* **31**, 14272 (2011).
- [23] B. B. Averbeck, P. E. Latham, and A. Pouget, *Neural Correlations, Population Coding and Computation*, *Nat. Rev. Neurosci.* **7**, 358 (2006).
- [24] R. Moreno-Bote, J. Beck, I. Kanitscheider, X. Pitkow, P. Latham, and A. Pouget, *Information-Limiting Correlations*, *Nat. Neurosci.* **17**, 1410 (2014).
- [25] P. Fries, *A Mechanism for Cognitive Dynamics: Neuronal Communication through Neuronal Coherence*, *Trends Cognit. Sci.* **9**, 474 (2005).
- [26] M. Abeles, *Corticonics: Neural Circuits of the Cerebral Cortex*, 1st ed. (Cambridge University Press, Cambridge, England, 1991).
- [27] M. Diesmann, M.-O. Gewaltig, and A. Aertsen, *Stable Propagation of Synchronous Spiking in Cortical Neural Networks*, *Nature (London)* **402**, 529 (1999).
- [28] N. Tabareau, J.-J. Slotine, and Q.-C. Pham, *How Synchronization Protects from Noise*, *PLoS Comput. Biol.* **6**, e1000637 (2010).
- [29] E. Salinas and T. J. Sejnowski, *Correlated Neuronal Activity and the Flow of Neural Information*, *Nat. Rev. Neurosci.* **2**, 539 (2001).
- [30] M. N. Shadlen and W. T. Newsome, *The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding*, *J. Neurosci.* **18**, 3870 (1998).
- [31] L. F. Abbott and P. Dayan, *The Effect of Correlated Variability on the Accuracy of a Population Code*, *Neural Comput.* **11**, 91 (1999).
- [32] B. Tripp and C. Eliasmith, *Neural Populations Can Induce Reliable Postsynaptic Currents without Observable Spike Rate Changes or Precise Spike Timing*, *Cereb. Cortex* **17**, 1830 (2007).
- [33] M. Schmuker and G. Schneider, *Processing and Classification of Chemical Data Inspired by Insect Olfaction*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 20285 (2007).
- [34] M. R. Cohen and J. H. R. Maunsell, *Attention Improves Performance Primarily by Reducing Interneuronal Correlations*, *Nat. Neurosci.* **12**, 1594 (2009).
- [35] T. Tetzlaff, A. Morrison, T. Geisel, and M. Diesmann, *Consequences of Realistic Network Size on the Stability of Embedded Synfire Chains*, *Neurocomputing: Variable Star Bulletin* **58-60**, 117 (2004).
- [36] B. Kriener, T. Tetzlaff, A. Aertsen, M. Diesmann, and S. Rotter, *Correlations and Population Dynamics in Cortical Networks*, *Neural Comput.* **20**, 2185 (2008).
- [37] T. Tetzlaff, S. Rotter, E. Stark, M. Abeles, A. Aertsen, and M. Diesmann, *Dependence of Neuronal Correlations on Filter Characteristics and Marginal Spike-Train Statistics*, *Neural Comput.* **20**, 2133 (2008).
- [38] A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, and N. K. Logothetis, *Decorrelated Neuronal Firing in Cortical Microcircuits*, *Science* **327**, 584 (2010).
- [39] C. Ly, J. Middleton, and B. Doiron, *Cellular and Circuit Mechanisms Maintain Low Spike Co-Variability and Enhance Population Coding in Somatosensory Cortex*, *Front. Comput. Neurosci.* **6**, 7 (2012).
- [40] J. Middleton, C. Omar, B. Doiron, and D. Simons, *Neural Correlation is Stimulus Modulated by Feedforward Inhibitory Circuitry*, *J. Neurosci.* **32**, 506 (2012).

- [41] M. T. Wiechert, B. Judkewitz, H. Riecke, and R. W. Friedrich, *Mechanisms of Pattern Decorrelation by Recurrent Neuronal Circuits*, *Nat. Neurosci.* **13**, 1003 (2010).
- [42] K. D. Harris and A. Thiele, *Cortical State and Attention*, *Nat. Rev. Neurosci.* **12**, 509 (2011).
- [43] J. S. Griffith and G. Horn, *An Analysis of Spontaneous Impulse Activity of Units in the Striate Cortex of Unrestrained Cats*, *J. Physiol.* **186**, 516 (1966).
- [44] K. W. Koch and J. M. Fuster, *Unit Activity in Monkey Parietal Cortex Related to Haptic Perception and Temporary Memory*, *Exp. Brain Res.* **76**, 292 (1989).
- [45] M. Shafi, Y. Zhou, J. Quintana, C. Chow, J. Fuster, and M. Bodner, *Variability in Neuronal Activity in Primate Cortex during Working Memory Tasks*, *Neuroscience (N.Y.)* **146**, 1082 (2007).
- [46] T. Hromadka, M. R. DeWeese, and A. M. Zador, *Sparse Representation of Sounds in the Unanesthetized Auditory Cortex*, *PLoS Biol.* **6**, e16 (2008).
- [47] D. H. O'Connor, S. P. Peron, D. Huber, and K. Svoboda, *Neural Activity in Barrel Cortex Underlying Vibrissa-Based Object Localization in Mice*, *Neuron* **67**, 1048 (2010).
- [48] A. Roxin, N. Brunel, D. Hansel, G. Mongillo, and C. van Vreeswijk, *On the Distribution of Firing Rates in Networks of Cortical Neurons*, *J. Neurosci.* **31**, 16217 (2011).
- [49] G. Buzsaki and K. Mizuseki, *The Log-Dynamic Brain: How Skewed Distributions Affect Network Operations*, *Nat. Rev. Neurosci.* **15**, 264 (2014).
- [50] S. Song, P. Sjöström, M. Reigl, S. Nelson, and D. Chklovskii, *Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits*, *PLoS Biol.* **3**, e68 (2005).
- [51] S. Lefort, C. Tamm, J.-C. F. Sarria, and C. C. H. Petersen, *The Excitatory Neuronal Network of the C2 Barrel Column in Mouse Primary Somatosensory Cortex*, *Neuron* **61**, 301 (2009).
- [52] A. A. Koulakov, T. Hromadka, and A. M. Zador, *Correlated Connectivity and the Distribution of Firing Rates in the Neocortex*, *J. Neurosci.* **29**, 3685 (2009).
- [53] M. Avermann, C. Tamm, C. Mateo, W. Gerstner, and C. Petersen, *Microcircuits of Excitatory and Inhibitory Neurons in Layer 2/3 of Mouse Barrel Cortex*, *J. Neurophysiol.* **107**, 3116 (2012).
- [54] Y. Ikegaya, T. Sasaki, D. Ishikawa, N. Honma, K. Tao, N. Takahashi, G. Minamisawa, S. Ujita, and N. Matsuki, *Interpyramid Spike Transmission Stabilizes the Sparseness of Recurrent Network Activity*, *Cereb. Cortex* **23**, 293 (2013).
- [55] A. Roxin, *The Role of Degree Distribution in Shaping the Dynamics in Networks of Sparsely Connected Spiking Neurons*, *Front. Comput. Neurosci.* **5**, 8 (2011).
- [56] A. Kuhn, A. Aertsen, and S. Rotter, *Neuronal Integration of Synaptic Input in the Fluctuation-Driven Regime*, *J. Neurosci.* **24**, 2345 (2004).
- [57] R. B. Stein, E. R. Gossen, and K. E. Jones, *Neuronal Variability: Noise or Part of the Signal?*, *Nat. Rev. Neurosci.* **6**, 389 (2005).
- [58] E. Marder and J.-M. Goaillard, *Variability, Compensation and Homeostasis in Neuron and Network Function*, *Nat. Rev. Neurosci.* **7**, 563 (2006).
- [59] C. Van Vreeswijk and H. Sompolinsky, *Chaotic Balanced State in a Model of Cortical Circuits*, *Neural Comput.* **10**, 1321 (1998).
- [60] M. Tsodyks, I. Mitkov, and H. Sompolinsky, *Pattern of Synchrony in Inhomogeneous Networks of Oscillators with Pulse Interactions*, *Phys. Rev. Lett.* **71**, 1280 (1993).
- [61] D. Golomb and J. Rinzel, *Dynamics of Globally Coupled Inhibitory Neurons with Heterogeneity*, *Phys. Rev. E* **48**, 4810 (1993).
- [62] L. Neltner, D. Hansel, G. Mato, and C. Meunier, *Synchrony in Heterogeneous Networks of Spiking Neurons*, *Neural Comput.* **12**, 1607 (2000).
- [63] M. Denker, M. Timme, M. Diesmann, F. Wolf, and T. Geisel, *Breaking Synchrony by Heterogeneity in Complex Networks*, *Phys. Rev. Lett.* **92**, 074103 (2004).
- [64] J. F. Mejias and A. Longtin, *Optimal Heterogeneity for Coding in Spiking Neural Networks*, *Phys. Rev. Lett.* **108**, 228102 (2012).
- [65] N. G. Stocks, *Suprathreshold Stochastic Resonance in Multilevel Threshold Systems*, *Phys. Rev. Lett.* **84**, 2310 (2000).
- [66] M. Shamir and H. Sompolinsky, *Implications of Neuronal Diversity on Population Coding*, *Neural Comput.* **18**, 1951 (2006).
- [67] M. I. Chelaru and V. Dragoi, *Efficient Coding in Heterogeneous Neuronal Populations*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 16344 (2008).
- [68] L. C. Osborne, S. E. Palmer, S. G. Lisberger, and W. Bialek, *The Neural Basis for Combinatorial Coding in a Cortical Population Response*, *J. Neurosci.* **28**, 13522 (2008).
- [69] K. Padmanabhan and N. N. Urban, *Intrinsic Biophysical Diversity Decorrelates Neuronal Firing While Increasing Information Content*, *Nat. Neurosci.* **13**, 1276 (2010).
- [70] G. Marsat and L. Maler, *Neural Heterogeneity and Efficient Population Codes for Communication Signals*, *J. Neurophysiol.* **104**, 2543 (2010).
- [71] L. A. Holmstrom, L. B. Eeuwes, P. D. Roberts, and C. V. Portfors, *Efficient Encoding of Vocalizations in the Auditory Midbrain*, *J. Neurosci.* **30**, 802 (2010).
- [72] M. Y. Yim, A. Aertsen, and S. Rotter, *Impact of Intrinsic Biophysical Diversity on the Activity of Spiking Neurons*, *Phys. Rev. E* **87**, 032710 (2013).
- [73] J. Lengler, F. Jug, and A. Steger, *Reliable Neuronal Systems: The Importance of Heterogeneity*, *PLoS One* **8**, e80694 (2013).
- [74] J. F. Mejias and A. Longtin, *Differential Effects of Excitatory and Inhibitory Heterogeneity on the Gain and Asynchronous State of Sparse Cortical Networks*, *Front. Comput. Neurosci.* **8**, 107 (2014).
- [75] S. J. Tripathy, K. Padmanabhan, R. C. Gerkin, and N. N. Urban, *Intermediate Intrinsic Diversity Enhances Neural Population Coding*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8248 (2013).
- [76] A. Bernacchia and X.-J. Wang, *Decorrelation by Recurrent Inhibition in Heterogeneous Neural Circuits*, *Neural Comput.* **25**, 1732 (2013).
- [77] J. Schemmel, A. Grübl, K. Meier, and E. Müller, in *Proceedings of the 2006 International Joint Conference*

- on *Neural Networks (IJCNN)*, Vancouver (IEEE, New York, 2006), pp. 1–6.
- [78] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevX.6.021023> for details on the hardware system, additional data-analysis and simulation results, as well as results obtained for a simplified linear model.
- [79] D. Badoni, M. Giulioni, V. Dante, and P. Del Giudice, in *Proceedings of the 2006 International Symposium on Circuits and Systems (ISCAS)*, Island of Kos (IEEE, New York, 2006), p. 4.
- [80] P. Hafliger, *Adaptive WTA with an Analog VLSI Neuromorphic Learning Chip*, *IEEE Trans. Neural Networks* **18**, 551 (2007).
- [81] R. Vogelstein, U. Mallik, J. Vogelstein, and G. Cauwenberghs, *Dynamically Reconfigurable Silicon Array of Spiking Neurons with Conductance-Based Synapses*, *IEEE Trans. Neural Networks* **18**, 253 (2007).
- [82] G. Indiveri, E. Chicca, and R. Douglas, *Artificial Cognitive Systems: From VLSI Networks of Spiking Neurons to Neuromorphic Cognition*, *Cogn. Comput.* **1**, 119 (2009).
- [83] R. Serrano-Gotarredona, M. Oster, P. Lichtsteiner, A. Linares-Barranco, R. Paz-Vicente, F. Gomez-Rodriguez, L. Camunas-Mesa, R. Berner, M. Rivas-Perez, T. Delbruck, S.-C. Liu, R. Douglas, P. Hafliger, G. Jimenez-Moreno, A. Ballcells, T. Serrano-Gotarredona, A. Acosta-Jimenez, and B. Linares-Barranco, *CAVIAR: A 45k Neuron, 5M Synapse, 12G Connects/s AER Hardware Sensory-Processing-Learning-Actuating System for High-Speed Visual Object Recognition and Tracking*, *IEEE Trans. Neural Networks* **20**, 1417 (2009).
- [84] S. Brink, S. Nease, P. Hasler, S. Ramakrishnan, R. Wunderlich, A. Basu, and B. Degnan, *A Learning-Enabled Neuron Array IC Based upon Transistor Channel Models of Biological Phenomena*, *IEEE Trans. Biomed. Circuits Syst.* **7**, 71 (2013).
- [85] B. Benjamin, P. Gao, E. McQuinn, S. Choudhary, A. Chandrasekaran, J. Bussat, R. Alvarez-Icaza, J. Arthur, P. Merolla, and K. Boahen, *Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations*, *Proc. IEEE* **102**, 699 (2014).
- [86] S. Renaud, J. Tomas, N. Lewis, Y. Bornat, A. Daouzli, M. Rudolph, A. Destexhe, and S. Saïghi, *PAX: A Mixed Hardware/Software Simulation Platform for Spiking Neural Networks*, *Neural Netw.* **23**, 905 (2010).
- [87] T. Yu and G. Cauwenberghs, *Analog VLSI Biophysical Neurons and Synapses with Programmable Membrane Channel Kinetics*, *IEEE Trans. Biomed. Circuits Syst.* **4**, 139 (2010).
- [88] A. Davison, D. Brüderle, J. M. Eppler, J. Kremkow, E. Müller, D. Pecevski, L. Perrinet, and P. Yger, *PyNN: A Common Interface for Neuronal Network Simulators*, *Front. Neuroinformatics* **2**, 11 (2009).
- [89] D. Brüderle, E. Müller, A. Davison, E. Müller, J. Schemmel, and K. Meier, *Establishing a Novel Modeling Tool: A Python-Based Interface for a Neuromorphic Hardware System*, *Front. Neuroinformatics* **3**, 17 (2009).
- [90] B. Kaplan, D. Brüderle, J. Schemmel, and K. Meier, in *Proceedings of the 2009 International Joint Conference on Neural Networks (IJCNN)*, Atlanta (IEEE, New York, 2009), pp. 1524–1530.
- [91] J. Bill, K. Schuch, D. Brüderle, J. Schemmel, W. Maass, and K. Meier, *Compensating Inhomogeneities of Neuromorphic VLSI Devices via Short-Term Synaptic Plasticity*, *Front. Comput. Neurosci.* **4**, 129 (2010).
- [92] D. Brüderle, J. Bill, B. Kaplan, J. Kremkow, K. Meier, E. Müller, and J. Schemmel, in *Proceedings of the 2010 International Symposium on Circuits and Systems (ISCAS)*, Paris (IEEE, New York, 2010).
- [93] Z. F. Mainen and T. J. Sejnowski, *Reliability of Spike Timing in Neocortical Neurons*, *Science* **268**, 1503 (1995).
- [94] J. De la Rocha, B. Doiron, E. Shea-Brown, J. Kresimir, and A. Reyes, *Correlation between Neural Spike Trains Increases with Firing Rate*, *Nature (London)* **448**, 802 (2007).
- [95] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, Cambridge, England, 2007).
- [96] A. A. Prinz, D. Bucher, and E. Marder, *Similar Network Activity from Disparate Circuit Parameters*, *Nat. Neurosci.* **7**, 1345 (2004).
- [97] T. Pfeil, A.-C. Scherzer, J. Schemmel, and K. Meier, in *Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN)* (IEEE, New York, 2013), pp. 1–5.
- [98] W. Bair, E. Zohary, and W. Newsome, *Correlated Firing in Macaque Visual Area MT: Time Scales and Relationship to Behavior*, *J. Neurosci.* **21**, 1676 (2001).
- [99] A. Kohn and M. A. Smith, *Stimulus Dependence of Neuronal Correlations in Primary Visual Cortex of the Macaque*, *J. Neurosci.* **25**, 3661 (2005).
- [100] R. Moreno-Bote and N. Parga, *Auto- and Crosscorrelations for the Spike Response of Leaky Integrate-and-Fire Neurons with Slow Synapses*, *Phys. Rev. Lett.* **96**, 028101 (2006).
- [101] M.-O. Gewaltig and M. Diesmann, *NEST (NEural Simulation Tool)*, *Scholarpedia* **2**, 1430 (2007).
- [102] W. Maass, *Noise as a Resource for Computation and Learning in Networks of Spiking Neurons*, *Proc. IEEE* **102**, 860 (2014).
- [103] M. London, A. Roth, L. Beeren, M. Häusser, and P. E. Latham, *Sensitivity to Perturbations In Vivo Implies High Noise and Suggests Rate Coding in Cortex*, *Nature (London)* **466**, 123 (2010).
- [104] D. Grytskyy, T. Tetzlaff, M. Diesmann, and M. Helias, *A Unified View on Weakly Correlated Recurrent Networks*, *Front. Comput. Neurosci.* **7**, 131 (2013).
- [105] M. Helias, T. Tetzlaff, and M. Diesmann, *Echoes in Correlated Neural Systems*, *New J. Phys.* **15**, 023002 (2013).
- [106] R. Rosenbaum and K. Josic, *Mechanisms that Modulate the Transfer of Spiking Correlations*, *Neural Comput.* **23**, 1261 (2011).
- [107] C. Ribault, K. Sekimoto, and A. Triller, *From the Stochasticity of Molecular Processes to the Variability of Synaptic Transmission*, *Nat. Rev. Neurosci.* **12**, 375 (2011).
- [108] T. O’Leary, A. H. Williams, A. Franci, and E. Marder, *Cell Types, Network Homeostasis, and Pathological*

- Compensation from a Biologically Plausible Ion Channel Expression Model*, *Neuron* **82**, 809 (2014).
- [109] M. Giulioni, P. Camilleri, M. Mattia, V. Dante, J. Braun, and P. Del Giudice, *Robust Working Memory in an Asynchronously Spiking Neural Network Realized in Neuromorphic VLSI*, *Front. Neurosci.* **5**, 149 (2012).
- [110] J. Schemmel, D. Brüderle, A. Grübl, M. Hock, K. Meier, and S. Millner, in *Proceedings of the 2010 International Symposium on Circuits and Systems (ISCAS)*, Paris (IEEE, New York, 2010), pp. 1947–1950.
- [111] E. Nordlie, M.-O. Gewaltig, and H. E. Plesser, *Towards Reproducible Descriptions of Neuronal Network Models*, *PLoS Comput. Biol.* **5**, e1000456 (2009).